ADVANCES IN MATHEMATICS 96, 226–263 (1992)

Hyperdeterminants

I. M. GELFAND

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903; and USSR Academy of Sciences

M. M. KAPRANOV

Department of Mathematics, Northwestern University, Evanston, Illinois 60208

AND

A. V. ZELEVINSKY

Department of Mathematics, Northeastern University, Boston, Massachusetts 02115

INTRODUCTION

The goal of this article is to provide a natural "higher-dimensional" generalization of the classical notion of the determinant of a square $n \times n$ matrix:

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sign}(\pi)\, a_{1, \pi(1)} \cdots a_{n, \pi(n)}. \tag{0.1}$$

By an r-dimensional "matrix" we shall mean an array $A = (a_{i_1, \dots, i_r})$ of numbers, where each index ranges over some finite set. We would like to extend the notion of the determinant to higher-dimensional matrices.

There were some attempts of a rather straightforward generalization of (0.1) for "hypercubic" matrices using summations over the product of several symmetric groups (see, e.g., [15, Sect. 54] and references therein).

In this paper we systematically develop another approach motivated by the theory of general discriminants worked out in our previous papers [6, 7]. The starting point of this approach is the observation that the variety of *degenerate* $n \times n$ matrices is projectively dual to the variety of matrices of rank 1. Now there is an obvious generalization of the rank 1 matrices for a higher-dimensional case: these are matrices of the form $(a_{i_1, \dots, i_r} = x^{(1)}_{i_1} x^{(2)}_{i_2} \cdots x^{(r)}_{i_r})$. Let X be the projectivization of the variety of these matrices: it is the product of r projective spaces in the classical Segre

0001-8708/92 $9.00

embedding. Following an analogy with a 2-dimensional case we say that an r-dimensional matrix $A = (a_{i_1, \ldots, i_r})$ is *degenerate* if its projectivization lies in the projective dual variety $X^\vee$. Here we identify the space of r-dimensional matrices of a given format with its dual by means of the pairing

$$(A, B) = \sum_{i_1, \ldots, i_r} a_{i_1, \ldots, i_r} b_{i_1, \ldots, i_r};$$

by definition, A is degenerate if and only if its orthogonal hyperplane is tangent to the variety of rank 1 matrices at some non-zero point.

It is easy to show that the variety $X^\vee$ is irreducible and defined over $\mathbf{Z}$. If it is a hypersurface then it is given by a unique (up to sign) irreducible polynomial over $\mathbf{Z}$ in matrix entries $a_{i_1, \ldots, i_r}$. We call this polynomial the *hyperdeterminant* and denote it $\operatorname{Det}(A)$. This is the main object of study in the present paper.

After arriving at the notion of a hyperdeterminant we discovered that this concept was originally introduced by A. Cayley [4] 150 years ago in essentially the same way. However, it should be mentioned that Cayley changed his terminology in later works and often meant by hyperdeterminant any invariant of a multilinear form.

The definition of the hyperdeterminant just given is quite implicit. It is not even clear from it for which matrix formats it does make sense, i.e., when $X^\vee$ is a hypersurface. In the case of ordinary (2-dimensional) matrices this happens exactly for square matrices. Thus matrix formats for which $X^\vee$ is a hypersurface can be thought of as higher-dimensional analogs of square matrices. To describe these formats it is convenient to keep the following convention. We will write the matrix format in the form $(l_1 + 1) \times \cdots \times (l_r + 1)$ and suppose that each matrix index i_k ranges over the set $[0, l_k] = \{0, 1, \ldots, l_k\}$, $k = 1, \ldots, r$. Then we have the following result.

THEOREM 0.1. *The hyperdeterminant of format* $(l_1 + 1) \times \cdots \times (l_r + 1)$ *exists if and only if*

$$l_k \leqslant \sum_{j \neq k} l_j \tag{0.2}$$

for all $k = 1, \ldots, r$.

In particular, for 2-dimensional matrices we obtain the usual condition $l_1 = l_2$.

Theorem 0.1 follows from more general results of F. Knop and G. Menzel [12] on the projective dual to the highest vector orbit in a representation of a reductive group. See Section 1 for details.

The hyperdeterminant has many interesting symmetries: we can group matrix indices in different ways and consider slices in various directions. In more invariant terms, an r-dimensional matrix represents a multilinear form $f \in V_1^* \otimes V_2^* \otimes \cdots \otimes V_r^*$, and the hyperdeterminant is invariant under the group $SL(V_1) \times \cdots \times SL(V_r)$, and also under permutations of the factors V_k^*. There are also some analogs of the multiplicative property $\det(AB) = \det(A)\det(B)$ of ordinary determinants. We discuss these basic properties in Section 1.

In fact, hyperdeterminants are particular case of A-discriminants (equations of projective duals of toric varieties) introduced and studied in [6]. Specializing the general machinery developed in [6] we express the hyperdeterminant of arbitrary format satisfying (0.2) as the determinant (or Whitehead torsion) of some complex of vector spaces (the Cayley–Koszul complex); see Section 2. The idea of representing multivariate discriminants and resultants as determinants of complexes is also due to Cayley, whose paper [3], published in 1848, introduces concepts now known as chain complexes, exact sequences, determinant (Whitehead torsion) of a based exact sequence, and Koszul complexes.

The machinery of Cayley–Koszul complexes leads in principle to an explicit formula for the hyperdeterminant as the ratio of two products of rather complicated determinants of some auxiliary matrices. It is not an easy task even to find the degree of the hyperdeterminant using this expression. But this can be done, and the degree can be expressed as the sum of certain positive summands of combinatorial significance (see Section 3 below). This expression is quite manageable and allows us to calculate the degree for any particular format without much trouble. It also allows us to give a new proof of Theorem 0.1.

The formula for the degree suggests the problem of finding a "purely polynomial" expression for the hyperdeterminant. This problem is still open in general, but we have solved it in some interesting special cases. The most important is the case of the so-called *boundary format* when one of the inequalities in (0.2) becomes an equality. This case is treated in detail in Section 4.

It is instructive to think of matrices of boundary format as proper higher-dimensional analogs of ordinary square matrices. The boundary hyperdeterminant can be expressed as an ordinary determinant very similar to the classical Sylvester determinantal expression for the resultant of two binary forms. In fact, the boundary hyperdeterminant can be identified with the resultant of a system of multilinear equations. It also has another interesting application: if we consider the r-dimensional format with $l_1 > l_2 + \cdots + l_r$ for which the variety $X^\vee$ has codimension greater than 1, then the Chow form of $X^\vee$ is given by the $(r+1)$-dimensional boundary hyperdeterminant of format $(l_1+1) \times \cdots \times (l_r+1) \times (l_1 - l_2 - \cdots - l_r + 1)$.

To the best of our knowledge, the only mathematician who ever studied hyperdeterminants after A. Cayley was L. Schläfli [18], in 1852. He developed an inductive (in r) approach to their calculation which worked in some cases but often produced not the hyperdeterminant itself but rather its product with some other polynomials ("fremde Faktoren"). In Section 5 we give a modern treatment of Schläfli's method and interpret his "fremde Faktoren" as Chow forms of components of the singular locus of the hyperdeterminantal variety.

This paper is only a first step toward the systematic study of hyperdeterminants, and we would be happy to attract the attention of mathematicians to this fascinating field. We would like to suggest three directions for further developments:

- Study of hyperdeterminantal varieties and their singularities.
- The combinatorial study of monomials in hyperdeterminants. Monomials in an ordinary determinant correspond to permutations. The monomials in general hyperdeterminants can be seen therefore as r-ary correspondences forming a "higher" analog of the symmetric group. Especially interesting are the "extreme" monomials, i.e., those corresponding to vertices of the Newton polytope [6].
- The generalization of eigenvalue theory to multidimensional matrices.

Speculating about possible applications of hyperdeterminants, one can mention algebraic equations of mathematical physics such as the Yang–Baxter equation and its generalizations where unknowns are multidimensional matrices [9]. One can hope that hyperdeterminants might be used in the study of their solutions.

1. Definition and Basic Properties of the Hyperdeterminant

Let $r \geqslant 2$ be an integer, and $A = (a_{i_1 \ldots i_r})$, $0 \leqslant i_k \leqslant l_k$, be an r-dimensional complex matrix of format $(l_1+1) \times \cdots \times (l_r+1)$.

The definition of the hyperdeterminant of A can be stated in geometric, analytic, or algebraic terms. Let us give all three formulations.

Geometrically, consider the product $X = P^{l_1} \times \cdots \times P^{l_r}$ of several projective spaces in the Segre embedding into the projective space $P^{(l_1+1)\cdots(l_r+1)-1}$ (if P^{l_k} is the projectivization of a vector space $V_k = \mathbf{C}^{l_k+1}$ then the ambient projective space is $P(V_1 \otimes \cdots \otimes V_r)$). Let $X^\vee$ be the *projective dual variety* of X consisting of all hyperplanes in $P^{(l_1+1)\cdots(l_r+1)-1}$ tangent to X at some point. Clearly, $X^\vee$ is an irreducible algebraic variety defined over $\mathbf{Q}$ in the dual projective space $(P^{(l_1+1)\cdots(l_r+1)-1})^*$. If $X^\vee$ is a hypersurface in $(P^{(l_1+1)\cdots(l_r+1)-1})^*$ then its defining equation, which is a homogeneous polynomial function on $V_1^* \otimes \cdots V_r^*$, is called the *hyperdeterminant of format* $(l_1+1)\times\cdots\times(l_r+1)$ and denoted Det. If each V_k is equipped by a basis then an element $f\in V_1^*\otimes\cdots\otimes V_r^*$ is represented by a matrix $A=(a_{i_1\cdots i_r})$, $0\leqslant i_k\leqslant l_k$, as above, and so $\mathrm{Det}(A)$ is a polynomial function of matrix entries. It is determined uniquely up to sign by the requirement that $\mathrm{Det}(A)$ have integer coefficients and be irreducible over $\mathbf{Z}$.

Analytically, the hyperplane given by the equation $f=0$ belongs to $X^\vee$ if and only if f vanishes at some point of X with all its first derivatives. If we choose the coordinate system $x^{(k)}=(x_0^{(k)}, x_1^{(k)}, ..., x_{l_k}^{(k)})$ on each V_k then an element $f\in V_1^*\otimes\cdots\otimes V_r^*$ is represented after restriction on X by a multilinear form

$$f(x^{(1)}, ..., x^{(r)})=\sum_{i_1, ..., i_r} a_{i_1\cdots i_r}x_{i_1}^{(1)}\cdots x_{i_r}^{(r)}. \tag{1.1}$$

Therefore, the condition $\mathrm{Det}(A)=0$ means that the system of equations

$$f(x)=\frac{\partial f(x)}{\partial x_i^{(k)}}=0 \tag{1.2}$$

(for all k, i) has a solution $x=(x^{(1)}, ..., x^{(r)})$ with all $x^{(k)}\neq 0$. We say that a multilinear form f or a matrix A satisfying this condition is *degenerate*.

The degeneracy of a form f can be characterized easily in terms of linear algebra. We denote by $\mathscr{K}(f)$ (or $\mathscr{K}(A)$) the set of points $x=(x^{(1)}, ..., x^{(r)})\in X=P^{l_1}\times\cdots\times P^{l_r}$ such that

$$f(x^{(1)}, ..., x^{(k-1)}, y, x^{(k+1)}, ..., x^{(r)})=0$$

for every $k=1, ..., r$ and $y\in V_k$. We shall sometimes call $\mathscr{K}(A)$ the *kernel* of A. For a bilinear form $f(x, y)$ there is a notion of left and right kernels

$$K_l(f)=\{x: f(x, y)=0, \forall y\}, \qquad K_r(f)=\{y: f(x, y)=0, \forall x\}$$

and $\mathscr{K}(f)=K_l(f)\times K_r(f)$.

PROPOSITION 1.1. *A form f is degenerate if and only if $\mathscr{K}(f)$ is nonempty.*

Proof. Computing the differential of f we see that $\mathscr{K}(f)$ is exactly the set of solutions of (1.2). ∎

In particular, if $r=2$ and so f is a bilinear form with a matrix A, the degeneracy of f just defined coincides with the usual notion of degeneracy and means that A is not of maximal rank. Obviously, this condition is of codimension 1 if and only if A is a square matrix, and in this case $\mathrm{Det}(A)$ coincides with the ordinary determinant $\det(A)$.

The following proposition can be found essentially in [18].

PROPOSITION 1.2. *Suppose that an r-dimensional matrix* $A_0 = (a_{i_1 \cdots i_r})$ *has a well-defined hyperdeterminant and is a smooth point of the hypersurface of degenerate matrices. Then* $\mathscr{X}(A_0)$ *consists of the unique point* $(x^{(1)}, ..., x^{(r)})$. *Furthermore, under a suitable normalization we have*

$$x^{(1)}_{i_1} \cdots x^{(r)}_{i_r} = \frac{\partial \operatorname{Det}(A)}{\partial a_{i_1 \cdots i_r}} \bigg|_{A = A_0} \tag{1.3}$$

for all $i_1, ..., i_r$.

Proof. This follows readily from the geometric definition of the hyperdeterminant and general properties of projective dual varieties. Namely, consider a projective variety $X \subset P(V)$ and its projective dual $X^\vee \subset P(V^*)$ (here V is a finite-dimensional complex vector space). Let $x \in X$, $\xi \in X^\vee$ be some smooth points. Then ξ can be thought of as a hyperplane in $P(V)$. The following "biduality theorem" is well known (see, e.g., [11]): the relation "ξ is tangent to X at x" is symmetric in ξ and x, i.e., coincides with the relation "x (considered as a hyperplane in $P(V^*)$) is tangent to $X^\vee$ at ξ."

In our situation $X = P^{l_1} \times \cdots \times P^{l_r}$ in the Segre embedding, so X is smooth. By our assumptions, $X^\vee$ is a hypersurface, and the matrix A_0 defines a smooth point ξ on this hypersurface. Then there is the unique $x \in X$ which is tangent to $X^\vee$ at ξ, and the homogeneous coordinates of x are the numbers $(\partial \operatorname{Det}(A)/\partial a_{i_1 \cdots i_r}|_{A = A_0})$. Combining this with the above general property we get our assertion. ∎

The first natural question about hyperdeterminants is to describe all matrix formats for which $\operatorname{Det}(A)$ is well-defined, i.e., $X^\vee$ is a hypersurface, or in other words, the degeneracy of A is a codimension 1 condition. The matrices of such formats can be viewed as multidimensional generalizations of ordinary square matrices.

THEOREM 1.3. *The hyperdeterminant of format* $(l_1 + 1) \times \cdots \times (l_r + 1)$ *exists if and only if*

$$l_k \leqslant \sum_{j \neq k} l_j \tag{1.4}$$

for all $k = 1, ..., r$.

Theorem 1.3 is a special case of a general result by F. Knop and G. Menzel [12]. To see this consider the following problem. Let G be a complex semisimple Lie group, and V an irreducible finite-dimensional

G-module with highest vector v. Let $X \subset P(V)$ be the projectivization of the G-orbit of v. The problem solved in [12] is to classify all pairs (G, V) such that the projective dual variety $X^\vee$ has codimension 1 in $P(V^*)$.

In order to apply this result to our situation we give one more interpretation of the hyperdeterminant, this time in representation-theoretic terms. Let $G = SL(V_1) \times \cdots \times SL(V_r)$; then $V = V_1 \otimes \cdots \otimes V_r$ is an irreducible G-module. It is easy to see that $X = P(V_1) \times \cdots \times P(V_r) \subset P(V)$ is a projectivization of the G-orbit of the highest vector of V. Now it is an easy exercise to deduce Theorem 1.3 from general criteria in [12].

In Section 3 we give another proof of Theorem 1.3, deducing it from the formula for the degree of Det.

Till the end of this section we assume that (1.4) holds, i.e., the hyperdeterminant of the matrix A exists. The next property of Det follows at once from any of the definitions.

Proposition 1.4. *The hyperdeterminant is relative invariant under the action of the group $GL(V_1) \times \cdots \times GL(V_r)$ (and so invariant under the action of $SL(V_1) \times \cdots \times SL(V_r)$).*

To state explicitly the consequences of Proposition 1.4 we need some terminology. We will identify the set of matrix (multi-)indices $I = \{(i_1, ..., i_r) : 0 \leqslant i_k \leqslant l_k\}$ of a matrix A with the set of vertices of the product $\Delta^{l_1} \times \cdots \times \Delta^{l_r}$ of r standard simplices. Thus, the submatrices of A correspond to faces of $\Delta^{l_1} \times \cdots \times \Delta^{l_r}$. By a *slice in the kth direction* we mean the subset of all indices in I with the fixed kth component, and also the corresponding submatrix of A. Two slices in the same direction are called *parallel.*

Corollary 1.5. (a) *Interchanging of two parallel slices leaves the hyperdeterminant invariant up to sign (which may be equal to 1).*

(b) *The hyperdeterminant is a homogeneous polynomial in the entries of each slice. The degree of homogeneity is the same for parallel slices.*

(c) *The hyperdeterminant does not change if we add to some slice a scalar multiple of a parallel slice.*

(d) *The hyperdeterminant of a matrix having two parallel slices proportional to each other is equal to 0. In particular,* $\mathrm{Det}(A) = 0$ *if A has a zero slice.*

Proof. The properties (a) to (c) express the (relative) invariance of $\mathrm{Det}(A)$ under the action of various elements from $GL(V_k) = GL(l_k + 1, \mathbf{C})$, namely, the permutation matrices, diagonal matrices, and unipotent matrices with only one non-zero off-diagonal entry, respectively. In fact, these matrices are known to generate the group $GL(l_k + 1, \mathbf{C})$, and so the combination of these three properties is equivalent to Proposition 1.4. The statement (d) follows at once from (b) and (c). ∎

Remark 1.6. It is clear that a polynomial $P(a_{i_1 \cdots i_r})$ satisfies the condition (c) from Corollary 1.5 if and only if P is annihilated by the differential operators

$$D_{ij}^{(1)} = \sum_{i_2, \ldots, i_r} a_{ji_2 \cdots i_r} \frac{\partial}{\partial a_{ii_2 \cdots i_r}} \qquad (i \neq j)$$

and similar operators for the slices of other directions. This is probably the most practical way of verifying this condition.

COROLLARY 1.7. *The degree of the hyperdeterminant of format* $(l_1+1) \times \cdots \times (l_r+1)$ *is a* (*not necessarily least*) *common multiple of the numbers* $l_1+1, \ldots, l_r+1$.

Proof. This follows at once from Corollary 1.5(b). ∎

Our next result will lead to a combinatorial characterization of the hyperdeterminant. This again requires some terminology. We define the *support* of a monomial $a^\alpha = \prod_{i_1, \ldots, i_r} a_{i_1 \cdots i_r}^{\alpha_{i_1 \cdots i_r}}$ as the set of all indices $(i_1, \ldots, i_r)$ such that $\alpha_{i_1 \cdots i_r} \neq 0$. By the *star* of an index $(i_1, \ldots, i_r)$ we mean the set of all indices which differ from $(i_1, \ldots, i_r)$ in at most one place. In other words, if we represent the indices by the vertices of the product of simplices then the star consists of a vertex itself and all the vertices connected to it by an edge. To give still another interpretation, we consider on the set $\prod_{k=1}^r [0, l_k]$ the *Hamming metric* used in the coding theory: the distance between two indices $(i_1, \ldots, i_r)$ and $(j_1, \ldots, j_r)$ is the number of positions k such that $i_k \neq j_k$. Then the star of an index $\mathbf{i} = (i_1, \ldots, i_r)$ is the Hamming ball of radius 1 with center at $\mathbf{i}$.

PROPOSITION 1.8. *For a polynomial* $P(a_{i_1 \cdots i_r})$ *the following conditions are equivalent*:

(a) *P is relative invariant w.r.t. the group* $GL(l_1+1, \mathbf{C}) \times \cdots \times GL(l_r+1, \mathbf{C})$ *and is divisible by* $\operatorname{Det}(A)$.

(b) *P satisfies the conditions* (a) *to* (c) *of Corollary* 1.5, *and there exists an index* $(i_1, \ldots, i_r)$ *such that the support of each monomial in* $P(a_{i_1 \cdots i_r})$ *meets the star of* $(i_1, \ldots, i_r)$.

(c) *P satisfies the conditions* (a) *to* (c) *of Corollary* 1.5, *and the support of each monomial in* $P(a_{i_1 \cdots i_r})$ *meets the star of every index* $(i_1, \ldots, i_r)$.

Proof. We mentioned in the proof of Corollary 1.5 that the relative invariance of P is equivalent to the fact that P satisfies the conditions (a) to (c) of Corollary 1.5. Suppose that P satisfies these conditions. By definition, P is divisible by $\mathrm{Det}(A)$ if and only if P vanishes at all degenerate matrices A. By Proposition 1.1, A is degenerate if and only if $\mathscr{K}(A) \neq \varnothing$. Using the invariance of P we can assume that $\mathscr{K}(A)$ contains a point $(x^{(1)}, ..., x^{(r)})$ such that each $x^{(k)}$ is a basis vector $e^{(k)}_{i_k}$. Then the condition $(x^{(1)}, ..., x^{(r)}) \in \mathscr{K}(A)$ means that $a_{j_1 \cdots j_r} = 0$ for all $(j_1, ..., j_r)$ in the star of $(i_1, ..., i_r)$. Clearly, P vanishes at all such matrices if and only if the support of each monomial in P meets the star of $(i_1, ..., i_r)$, and we are done. ∎

Note that the second assertion in part (b) means that the support of any monomial from Det forms a 1-net in $\prod_{k=1}^{r} [0, l_k]$ with respect to the Hamming metric. Such nets are known as error-correcting codes. It would be interesting to study monomials from this point of view.

To illustrate the use of Proposition 1.8 we will give an explicit formula for the hyperdeterminant of the minimal 3-dimensional format $2 \times 2 \times 2$. The hyperdeterminant in this case was already known to A. Cayley (see [4, p. 89]).

PROPOSITION 1.9. *The hyperdeterminant of the matrix* $A = (a_{ijk})$ $(i, j, k = 0, 1)$ *is given by the formula*

$$\begin{aligned}\mathrm{Det}(A) = {} & (a_{000}^2 a_{111}^2 + a_{001}^2 a_{110}^2 + a_{010}^2 a_{101}^2 + a_{011}^2 a_{100}^2) \\ & - 2(a_{000} a_{001} a_{110} a_{111} + a_{000} a_{010} a_{101} a_{111} + a_{000} a_{011} a_{100} a_{111} \\ & + a_{001} a_{010} a_{101} a_{110} + a_{001} a_{011} a_{110} a_{100} + a_{010} a_{011} a_{101} a_{100}) \\ & + 4(a_{000} a_{011} a_{101} a_{110} + a_{001} a_{010} a_{100} a_{111}). \end{aligned} \tag{1.5}$$

Proof. We claim that the polynomial P defined by (1.5) satisfies conditions of Proposition 1.8. The properties (a) to (c) are verified directly (to verify (c) one can use Remark 1.6).

The monomials appearing in P can be visualized as follows. If we represent matrix entries by the vertices of the cube then monomials in the first group correspond to four main diagonals of the cube, monomials in the second group to six rectangles formed by pairs of opposite edges, and monomials in the third group to two tetrahedra whose edges are diagonals of the cube's faces. Obviously, the support of each of these monomials meets the star of every vertex of the cube. This shows that P satisfies conditions of Proposition 1.8 and hence is divisible by $\mathrm{Det}(A)$.

There are several ways to complete the proof, i.e., to show that $P = \mathrm{Det}(A)$. For instance, it is not hard to see that P is irreducible. One can also refer to the results of Section 3, which show that $\mathrm{Det}(A)$ in our case has degree 4 (see, e.g., Corollary 3.9 below). A more elementary possibility is to use Corollary 1.7, which in our case means that $\mathrm{Det}(A)$ has even degree, and observe that there are no polynomials of degree 2 satisfying conditions of Proposition 1.8. In fact, one can show that in our case $\mathrm{Det}(A)$ is the $SL(2)\times SL(2)\times SL(2)$ invariant of minimal degree. ▮

Note that by definition, vanishing of $\mathrm{Det}(A)$ for a $2\times 2\times 2$ matrix A (or, equivalently, degeneracy of A) means that the following system of six homogeneous equations with six unknowns has a non-trivial solution:

$$\begin{aligned}
a_{000}x_0y_0+a_{010}x_0y_1+a_{100}x_1y_0+a_{110}x_1y_1&=0\\
a_{001}x_0y_0+a_{011}x_0y_1+a_{101}x_1y_0+a_{111}x_1y_1&=0\\
a_{000}x_0z_0+a_{001}x_0z_1+a_{100}x_1z_0+a_{101}x_1z_1&=0\\
a_{010}x_0z_0+a_{011}x_0z_1+a_{110}x_1z_0+a_{111}x_1z_1&=0\\
a_{000}y_0z_0+a_{001}y_0z_1+a_{010}y_1z_0+a_{011}y_1z_1&=0\\
a_{100}y_0z_0+a_{101}y_0z_1+a_{110}y_1z_0+a_{111}y_1z_1&=0.
\end{aligned}\tag{1.6}$$

It is not so easy (although possible) to prove directly that the system (1.6) has a non-trivial solution if and only if the expression (1.5) vanishes.

Now we give a multidimensional generalization of the fact that the determinant is preserved by transposition of a matrix. For a matrix $A=(a_{i_1\cdots i_r})$ of format $(l_1+1)\times\cdots\times(l_r+1)$ and a permutation σ of indices $1,\ldots,r$ we denote by $\sigma(A)$ the matrix of format $(l_{\sigma^{-1}(1)}+1)\times\cdots\times(l_{\sigma^{-1}(r)}+1)$, whose $(j_1,j_2,\ldots,j_r)$th entry is equal to $a_{j_{\sigma(1)}\cdots j_{\sigma(r)}}$. The following result is an immediate consequence of definitions.

Proposition 1.10, *If A is degenerate then $\sigma(A)$ is degenerate for every permutation σ. If* $\mathrm{Det}(A)$ *exists then* $\mathrm{Det}(\sigma(A))$ *also exists and it is equal to* $\mathrm{Det}(A)$.

Now we discuss an analog of the multiplicative property of the ordinary determinant. Let $A=(a_{i_1\cdots i_r})$ be a matrix of format $(l_1+1)\times\cdots\times(l_r+1)$ and $B=(b_{j_1\cdots j_s})$ be a matrix of format $(m_1+1)\times\cdots\times(m_s+1)$. Suppose that $l_r=m_1$. We define the convolution (or product) $A*B$ of A and B to be the $(r+s-1)$-dimensional matrix C of format $(l_1+1)\times\cdots\times(l_{r-1}+1)\times(m_2+1)\times\cdots\times(m_s+1)$ with entries

$$c_{i_1\cdots i_{r-1}j_2\cdots j_s}=\sum_{h=0}^{l_r}a_{i_1\cdots i_{r-1}h}b_{hj_2\cdots j_s}.$$

Similarly, we can define the convolution $A*_{p,q}B$ with respect to a pair of indices p, q such that $l_p=m_q$.

PROPOSITION 1.11. *If A, B are degenerate then $A * B$ is also degenerate.*

Proof. We shall use Proposition 1.1. The definitions readily imply that if $(x^{(1)}, ..., x^{(r)}) \in \mathscr{K}(A)$ and $(y^{(1)}, ..., y^{(s)}) \in \mathscr{K}(B)$ then $(x^{(1)}, ..., x^{(r-1)}, y^{(2)}, ..., y^{(s)}) \in \mathscr{K}(A * B)$. ∎

COROLLARY 1.12. *Let formats of A, B be such that hyperdeterminants of A, B, and $A * B$ are defined. Then there exist polynomials $P(A, B)$ and $Q(A, B)$ in entries of A and B such that*

$$\mathrm{Det}(A * B) = P(A, B)\,\mathrm{Det}(A) + Q(A, B)\,\mathrm{Det}(B). \tag{1.7}$$

Since $\mathrm{Det}(A)$ and $\mathrm{Det}(B)$ depend on disjoint sets of variables, it follows that P and Q in (1.7) are defined uniquely up to transformations

$$P \mapsto P + R(A, B)\,\mathrm{Det}(B), \qquad Q \mapsto Q - R(A, B)\,\mathrm{Det}(A).$$

It would be interesting to study multiplicative properties of hyperdeterminants in more detail.

2. CAYLEY–KOSZUL COMPLEXES

In this section we associate to an r-dimensional matrix A a family of complexes of finite-dimensional vector spaces which we call *Cayley–Koszul complexes*. We will show that the hyperdeterminant $\mathrm{Det}(A)$ can be expressed as the determinant (or the Whitehead torsion) of an appropriate Cayley–Koszul complex.

As in Section 1, we interpret A as the matrix of a multilinear form f on the product $V_1 \times \cdots \times V_r$, where each V_k is a finite-dimensional vector space with coordinates $x_0^{(k)}, x_1^{(k)}, ..., x_{l_k}^{(k)}$.

Let Ω^p be the space of polynomial differential p-forms on $V_1 \times \cdots \times V_r$. For each vector field ξ on $V_1 \times \cdots \times V_r$ let $\iota_\xi : \Omega^p \to \Omega^{p-1}$ be the inner derivative, and $L_\xi = d\iota_\xi + \iota_\xi d$ the Lie derivative. For $k = 1, ..., r$ let

$$\xi_k = \sum_i x_i^{(k)} \frac{\partial}{\partial x_i^{(k)}}$$

be the Euler field on V_k considered as a vector field on $V_1 \times \cdots \times V_r$.

Now let $m_1, ..., m_r \in \mathbf{Z}$. We will associate to $m_1, ..., m_r$ and f the Cayley–Koszul complex $C^\cdot = C^\cdot(m_1, ..., m_r; f) : \{C^0 \xrightarrow{\partial_f} C^1 \xrightarrow{\partial_f} \cdots\}$. We define $C^p = C^p(m_1, ..., m_r)$ to be the space of all differential p-forms ω with polynomial coefficients on $V_1 \times \cdots \times V_r$ satisfying conditions

$$L_{\xi_k}(\omega) = (p + m_k)\omega \qquad (k = 1, ..., r) \tag{2.1}$$

and

$$\iota_{\xi_1}(\omega) = \iota_{\xi_2}(\omega) = \cdots = \iota_{\xi_r}(\omega). \tag{2.2}$$

The differential $\partial_f: C^p \to C^{p+1}$ is defined as the exterior multiplication by the 1-form df.

It is easy to check that the Cayley–Koszul complex is well-defined. Its terms do not depend on f. In fact, the condition (2.1) means simply that $\omega \in C^p(m_1, ..., m_r)$ has the degree of homogeneity $(p+m_k)$ with respect to the variables $x_0^{(k)}, x_1^{(k)}, ..., x_{l_k}^{(k)}$, where we assume that $\deg(x_i^{(k)}) = \deg(dx_i^{(k)}) = 1$.

THEOREM 2.1. *Suppose all m_k are nonnegative integers. Then the Cayley–Koszul complex $C^\cdot(m_1, ..., m_r; f)$ fails to be exact if and only if f is degenerate.*

Proof. Let us give a sheaf-theoretic interpretation of the Cayley–Koszul complexes. For $m_1, ..., m_r \in \mathbf{Z}$ let $\mathcal{O}(m_1, ..., m_r)$ be the sheaf on $X = P^{l_1} \times \cdots \times P^{l_r}$ whose sections are homogeneous functions of degree m_k in coordinates $(x_0^{(k)} : x_1^{(k)} : \cdots : x_{l_k}^{(k)})$ on each P^{l_k}. We will identify a matrix A or a form f as above with a global section of the sheaf $\mathcal{O}(1, 1, ..., 1)$. Let $\mathcal{T} = J^1(\mathcal{O}(1, 1, ..., 1))$ be the first jet bundle of $\mathcal{O}(1, 1, ..., 1)$; cf. [17]. To every (local) section s of $\mathcal{O}(1, 1, ..., 1)$ a section $j(s)$ of $\mathcal{T}$ is associated.

Now consider the Koszul complex of sheaves

$$\mathscr{C}^\cdot(f): \{\mathcal{O} \xrightarrow{\partial_f} \mathcal{T} \xrightarrow{\partial_f} \Lambda^2(\mathcal{T}) \cdots \xrightarrow{\partial_f} \Lambda^{l_1 + \cdots + l_r + 1}(\mathcal{T})\}$$

on X, where the differential ∂_f is the exterior multiplication by $j(f)$.

LEMMA 2.2. *For all $m_1, ..., m_r \in \mathbf{Z}$ the Cayley–Koszul complex $C^\cdot(m_1, ..., m_r; f)$ is naturally isomorphic to the complex of global sections of the twisted Koszul complex $\mathscr{C}^\cdot(f) \otimes \mathcal{O}(m_1, ..., m_r)$.*

Proof. First consider the case $p = 1$, $m_1 = \cdots = m_r = 0$. For any open set $U \subset X$ let $\tilde{U}$ be its inverse image in $(V_1 - 0) \times \cdots \times (V_r - 0)$. Then we can identify sections of $\mathcal{T}$ over U with 1-forms on $\tilde{U}$ satisfying (2.1) and (2.2). This identification is given by the mapping $j(f) \mapsto df$. Passing to exterior powers and tensoring with $\mathcal{O}(m_1, ..., m_r)$, we can now define the mapping from $H^0(U, \Lambda^p(\mathcal{T}) \otimes \mathcal{O}(m_1, ..., m_r))$ to the space of p-forms on $\tilde{U}$ satisfying (2.1) and (2.2). This is in fact an isomorphism of sheaves over X, which can be checked easily in local coordinates. Taking global sections, we obtain the desired result. ∎

LEMMA 2.3. *A form f is degenerate if and only if the Koszul complex of sheaves $\mathscr{C}^\cdot(f)$ fails to be exact.*

Proof. A form f is degenerate if and only if the section $j(f) \in H^0(X, J^1(\mathcal{O}(1, 1, ..., 1))$ vanishes at some point (this is simply a reformulation of conditions (1.2)). But this is exactly when the Koszul complex of sheaves fails to be exact; see, e.g., [8, Chap. 5, Sect. 3]. ∎

Clearly, the tensor multiplication by the invertible sheaf $\mathcal{O}(m_1, ..., m_r)$ does not affect the property of a complex of sheaves to be exact. By the general property (see [8, Chap. 3, Sect. 5]), the functor of taking global sections preserves exactness if all the terms of a complex of sheaves do not have higher cohomology. Therefore, our theorem becomes a consequence of the following lemma.

LEMMA 2.4. *If all m_k are nonnegative integers then $H^i(X, \Lambda^p(\mathcal{T}) \otimes \mathcal{O}(m_1, ..., m_r)) = 0$ for all p and all $i > 0$.*

Proof. Let Ω^p_X be the sheaf of differential p-forms on X, and $\Omega^p(m_1, ..., m_r)$ be the sheaf $\Omega^p_X \otimes \mathcal{O}(m_1, ..., m_r)$. By definition of the jet bundle, we have an exact sequence of vector bundles over X

$$0 \to \Omega^1(1, ..., 1) \to \mathcal{T} \to \mathcal{O}(1, ..., 1) \to 0. \tag{2.3}$$

Passing in (2.3) to exterior powers and tensoring with $\mathcal{O}(m_1, ..., m_r)$ we obtain for each p the exact sequence

$$\begin{aligned} 0 &\to \Omega^p(p+m_1, ..., p+m_r) \to \Lambda^p(\mathcal{T}) \otimes \mathcal{O}(m_1, ..., m_r) \\ &\to \Omega^{p-1}(p+m_1, ..., p+m_r) \to 0. \end{aligned} \tag{2.4}$$

Therefore, it is enough to prove that $H^i(X, \Omega^p(p+m_1, ..., p+m_r)) = 0$ for all p and all $i > 0$.

The sheaf $\Omega^p(p+m_1, ..., p+m_r)$ on X can be decomposed as

$$\begin{aligned} &\Omega^p(p+m_1, ..., p+m_r) \\ &\quad = \bigoplus_{p_1 + \cdots + p_r = p} (\pi_1^* \Omega^{p_1}(p+m_1) \otimes \cdots \otimes \pi_r^* \Omega^{p_r}(p+m_r)), \end{aligned} \tag{2.5}$$

where each $\Omega^{p_k}(p+m_k)$ is a sheaf on P^{l_k}, and π_k is the projection of $X = P^{l_1} \times \cdots \times P^{l_r}$ onto the kth factor. By the Künneth formula, we have

$$\begin{aligned} &H^i(X, \Omega^p(p+m_1, ..., p+m_r)) \\ &\quad = \bigoplus_{i_1 + \cdots + i_r = i} \; \bigoplus_{p_1 + \cdots + p_r = p} (H^{i_1}(P^{l_1}, \Omega^{p_1}(p+m_1)) \otimes \cdots \\ &\qquad \otimes H^{i_r}(P^{l_r}, \Omega^{p_r}(p+m_r))). \end{aligned}$$

So our statement follows at once from the next lemma, which is a special case of Bott's theorem (see, e.g., [14, Chap. 1, Sect. 1]).

LEMMA 2.5. *The cohomology $H^i(P^l, \Omega^p(m))$ vanishes whenever $i>0$ and $m \geqslant p \geqslant 0$.*

Lemma 2.4, and hence Theorem 2.1, is proven. ∎

The group $GL(V_1)\times \cdots \times GL(V_r)$ acts naturally on the Cayley–Koszul complex. To describe this action we need the following notation. For any finite-dimensional vector space V denote by $S^{(i\mid j)}(V)$ the irreducible $GL(V)$-module corresponding to the *hook* Young diagram having one row of length i and j rows of length 1 (i.e., one column of length $j+1$). We will use the following realization of this module.

LEMMA 2.6. *For any finite-dimensional vector space V the $GL(V)$-module $S^{(i\mid j)}(V^*)$ is isomorphic to the space $B^i(j)$ of polynomial differential i-forms ω on V such that $L_\xi(\omega)=(i+j)\omega$, $\iota_\xi(\omega)=0$, where ξ is the Euler vector field on V.*

Proof. By definition, $S^{(i\mid j)}(V)$ is the irreducible $GL(V)$-module with highest weight $(i, 1, ..., 1, 0, ..., 0)$ (with j units) and $S^{(i\mid j)}(V^*)$ is its dual. It is easy to see that $B^i(j)$ has a unique highest vector, and the weight of this vector is the same as that in $S^{(i\mid j)}(V^*)$.

PROPOSITION 2.7. *Each term $C^p(m_1, ..., m_r)$ of the Cayley–Koszul complex is a multiplicity free module over $GL(V_1)\times \cdots \times GL(V_r)$ isomorphic to*

$$\bigoplus_{p_1, ..., p_r} (S^{(p+m_1-p_1\mid p_1)}(V_1)\otimes \cdots \otimes S^{(p+m_r-p_r\mid p_r)}(V_r)),$$

the sum over all $p_1, ..., p_r$ with $p_1+\cdots+p_r=p$ or $p-1$.

Proof. Let $B^p(m_1, ..., m_r)$ denote the subspace of $C^p(m_1, ..., m_r)$ consisting of p-forms ω satisfying conditions (2.1) and

$$\iota_{\xi_1}(\omega)=\iota_{\xi_2}(\omega)=\cdots=\iota_{\xi_r}(\omega)=0. \tag{2.2'}$$

By definition, $B^p(m_1, ..., m_r)$ can be identified with the space of global sections of the sheaf $\Omega^p(p+m_1, ..., p+m_r)$ on X. Clearly, $B^p(m_1, ..., m_r)$ is the kernel of the projection $C^p(m_1, ..., m_r)\to B^{p-1}(m_1+1, ..., m_r+1)$ sending each $\omega \in C^p(m_1, ..., m_r)$ to $\omega'=\iota_{\xi_1}(\omega)=\iota_{\xi_2}(\omega)=\cdots=\iota_{\xi_r}(\omega)$. Therefore, $C^p(m_1, ..., m_r)$ is isomorphic as a $GL(V_1)\times\cdots\times GL(V_r)$-module to $B^p(m_1, ..., m_r)\oplus B^{p-1}(m_1+1, ..., m_r+1)$.

Taking the global sections of both parts of (2.5), we obtain the decomposition

$$B^p(m_1, ..., m_r)$$
$$= \bigoplus_{p_1 + \cdots + p_r = p} (B^{p_1}(p + m_1 - p_1) \otimes \cdots \otimes B^{p_r}(p + m_r - p_r)), \tag{2.6}$$

and the analogous decomposition for $B^{p-1}(m_1 + 1, ..., m_r + 1)$. It remains to apply Lemma 2.6. ▌

Theorem 2.1 allows us to give an explicit formula for $\mathrm{Det}(A)$ as the *determinant* of the Cayley–Koszul complex. For the convenience of the reader we recall the definition of the determinant of a complex [2, 3]. Suppose we are given an exact sequence $(C^{\cdot}, d)$ of finite-dimensional vector space over some field k:

$$(C^{\cdot}, d) = \{C^0 \xrightarrow{d} C^1 \xrightarrow{d} \cdots \xrightarrow{d} C^m\}.$$

Suppose also that a basis B_p is chosen in each term C^p, so the differential $d: C^p \to C^{p+1}$ is represented by a matrix D_p whose columns are indexed by B_p, and rows are indexed by B_{p+1}. To these data one can associate a nonzero element $\det(C^{\cdot}, d, (B_p)) \in k$. If our complex has the form

$$\{0 \longrightarrow C^1 \xrightarrow{d} C^2 \longrightarrow 0\},$$

where d is an isomorphism with the matrix D, then $\det(C^{\cdot}, d, (B_p))$ is the usual determinant of D. In general it is given by the following construction.

For any subsets $J \subset B_p$, $I \subset B_{p+1}$ let $D_p(I, J)$ denote the submatrix of D_p with rows from I and columns from J. We say that a collection $\{I_p \subset B_p\}$ is *admissible* if for any $p = 1, ..., m-1$ we have $\mathrm{Card}(B_p - I_p) = \mathrm{Card}(I_{p+1})$, and the submatrix $D_p(I_{p+1}, B_p - I_p)$ is invertible. If $\{I_p\}$ is an admissible collection then we define

$$\det(C^{\cdot}, d, (B_p)) = \prod_p \det(D_p(I_{p+1}, B_p - I_p))^{(-1)^p} \tag{2.7}$$

(one can show that admissible collections exist for any exact sequence, and that the RHS of (2.7) does not depend on the choice of an admissible collection; see, e.g., [2]).

We shall apply the determinant construction to Cayley–Koszul complexes. Fix nonnegative integers $m_1, ..., m_r$, and choose some basis B_p in the vector space $C^p(m_1, ..., m_r)$ for $p = 0, 1, ..., l_1 + \cdots + l_r + 1$. The correspondence $f \mapsto \det(C^{\cdot}(m_1, ..., m_r), \partial_f, (B_p))$ defines a rational function of f.

A different choice of the system of bases B_p results in multiplying this function by a non-zero constant.

Let A be a multidimensional matrix corresponding to a form f. In the next theorem we use the convention that $\mathrm{Det}(A)=1$ if the subvariety of degenerate matrices has codimension greater than 1.

THEOREM 2.8. *For all non-negative integers $m_1, ..., m_r$ we have*

$$\mathrm{Det}(A)=\det(C^{\cdot}(m_1, ..., m_r), \partial_f, (B_p))^{(-1)^{l_r+\cdots l_r+1}} \tag{2.8}$$

for some choice of bases B_p.

If all m_k are equal to each other then Theorem 2.8 is a special case of Corollary 2D.11 in [6]. The same arguments apply to the general case.

Note that the formula (2.8) leads to a quite complicated expression of the hyperdeterminant. In Section 4 we consider an important special case of the so-called *boundary format* when a much simpler formula is available.

COROLLARY 2.9. *The degree of the hyperdeterminant is equal to*

$$\sum_{p=0}^{l_1+\cdots+l_r+1} (-1)^{l_1+\cdots+l_r+1-p}\, p \dim C^p(m_1, ..., m_r), \tag{2.9}$$

independently of the choice of nonnegative m_k.

This follows immediately from (2.7).

3. DEGREE OF THE HYPERDETERMINANT

Fix $r\geqslant 2$ and let $N(l_1, ..., l_r)$ be the degree of the hyperdeterminant of format $(l_1+1)\times\cdots\times(l_r+1)$ (if this hyperdeterminant is not defined we make the convention $N(l_1, ..., l_r)=0$). The following theorem was obtained by the authors together with D. R. Leshchiner.

THEOREM 3.1. *The generating function for the degree $N(l_1, ..., l_r)$ is given by*

$$\sum_{l_1, ..., l_r} N(l_1, ..., l_r)\, z_1^{l_1}\cdots z_r^{l_r}=\left(1-\sum_{k=2}^{r}(k-1)\, e_k(z_1, ..., z_r)\right)^{-2}, \tag{3.1}$$

where $e_k(z_1, ..., z_r)$ is the kth elementary symmetric polynomial.

Proof. We will use the general formula for the degree of the $\mathcal{A}$-discriminant given in [6]. Let us recall that to each finite subset $\mathcal{A}$ of an integer lattice $\mathbf{Z}^n$ is associated a toric variety $X_{\mathcal{A}}$. Let $Q\subset\mathbf{R}^n$ be the convex hull of $\mathscr{A}$. The following proposition was proven in [6, Corollary 2C.6(b).].

PROPOSITION 3.2. *If the variety $X_{\mathscr{A}}$ is smooth then the degree of its projective dual is given by*

$$\deg(X_{\mathscr{A}}^{\vee}) = \sum_{\Gamma \subset Q} (-1)^{\dim(Q) - \dim(\Gamma)} (\dim(\Gamma) + 1) \operatorname{Vol}(\Gamma), \tag{3.2}$$

where the sum is over all faces $\Gamma \subset Q$, and the volume form Vol *on each face Γ is normalized so that an elementary simplex on the lattice affinely spanned by $\mathscr{A} \cap \Gamma$ has volume* 1.

Now note that the variety $X = P^{l_1} \times \cdots \times P^{l_r}$ in the Segre embedding is of the form $X_{\mathscr{A}}$, where $\mathscr{A}$ is the set of vertices of the product $Q = \Delta^{l_1} \times \cdots \times \Delta^{l_r}$ of several standard simplices. It follows that the degree $N(l_1, ..., l_r)$ is given by the formula (3.2) for this particular convex polytope Q. In this case the formula (3.2) can be deduced from (2.9) by some limit procedure.

Clearly, each face $\Gamma \subset Q$ has the form $\Gamma = \Delta^{m_1} \times \cdots \times \Delta^{m_r}$ for some $0 \leqslant m_k \leqslant l_k$, and for given $m_1, ..., m_r$ there are $\prod_k \binom{l_k+1}{m_k+1}$ faces of this type. Furthermore, for a face $\Gamma = \Delta^{m_1} \times \cdots \times \Delta^{m_r}$ we have

$$\dim(\Gamma) = m_1 + \cdots + m_r, \quad \operatorname{Vol}(\Gamma) = \binom{m_1 + \cdots + m_r}{m_1, ..., m_r}.$$

Substituting all this into (3.2) we get

$$N(l_1, ..., l_r) = \sum_{0 \leqslant m_k \leqslant l_k} (-1)^{\sum (l_k - m_k)} (m_1 + \cdots + m_r + 1) \times \binom{m_1 + \cdots + m_r}{m_1, ..., m_r} \prod_k \binom{l_k + 1}{m_k + 1}.$$

Letting $l_k - m_k = p_k$ we see that our generating function takes the form

$$\sum_{l_1, ..., l_r} N(l_1, ..., l_r)\, z_1^{l_1} \cdots z_r^{l_r} = \sum_{m_1, ..., m_r} (m_1 + \cdots + m_r + 1) \binom{m_1 + \cdots + m_r}{m_1, ..., m_r} z_1^{m_1} \cdots z_r^{m_r} \times \prod_k \sum_{p_k} (-1)^{p_k} \binom{m_k + p_k + 1}{m_k + 1} z_k^{p_k}. \tag{3.3}$$

By the binomial formula, the inner sum in the RHS of (3.3) is equal to $(1+z_k)^{-m_k-2}$. It follows that our generating function is equal to

$$\prod_k (1+z_k)^{-2} \sum_{s \geqslant 0} (s+1) \sum_{m_1+\cdots+m_r=s} \binom{s}{m_1, ..., m_r} \prod_k \left(\frac{z_k}{1+z_k}\right)^{m_k}. \quad (3.4)$$

By the multinomial formula, the inner sum in (3.4) is equal to $(\sum_k (z_k/(1+z_k)))^s$, and then the summation over s gives us $(1-\sum_k (z_k/(1+z_k)))^{-2}$. Therefore, we get

$$\sum_{l_1, ..., l_r} N(l_1, ..., l_r)\, z_1^{l_1} \cdots z_r^{l_r} = \left[\prod_k (1+z_k) - \sum_k z_k \prod_{j \neq k} (1+z_j)\right]^{-2}. \quad (3.5)$$

It remains to observe that the polynomial $[\prod_k (1+z_k) - \sum_k z_k \prod_{j\neq k} (1+z_j)]$ contains only square-free monomials in $z_1, ..., z_r$, and every such monomial of degree k occurs with the coefficient $(1-k)$. Theorem 3.1 is proven. ▮

We will deduce from (3.1) the combinatorial formula for $N(l_1, ..., l_r)$ expressing this degree as the sum of positive summands of combinatorial significance. To do this we need some basic facts about partitions and symmetric polynomials. We will follow the terminology of [13]. By a *partition* we mean a finite sequence $\lambda = (l_1, ..., l_r)$ of nonnegative integers defined up to rearranging of the terms and adding or deleting an arbitrary number of zeros. The positive terms l_k are called *parts* of λ. The number of parts equal to i is denoted by $m_i = m_i(\lambda)$, and λ is also written as $\lambda = (1^{m_1}, 2^{m_2}, ...)$.

For every two partitions $\lambda = (l_1, ..., l_r)$ and $\nu = (n_1, ..., n_s)$ let $M_{\lambda\nu}$ denote the number of (0, 1)-matrices of format $r \times s$ with row sums $l_1, ..., l_r$ and column sums $n_1, ..., n_s$.

Corollary 3.3. *For every partition $\lambda = (l_1, ..., l_r)$ we have*

$$N(l_1, ..., l_r) = \sum_\nu (m_2 + \cdots + m_p + 1)! \left(\prod_{i=2}^{p} \frac{(i-1)^{m_i}}{m_i!}\right) \cdot M_{\lambda\nu}, \quad (3.6)$$

the sum over all partitions $\nu = (1^{m_1}, 2^{m_2}, ..., p^{m_p})$ with $m_1 = 0$.

Proof. For $\nu = (1^{m_1}, 2^{m_2}, ..., p^{m_p})$ we put $e_\nu(z_1, ..., z_r) = \prod_{i=1}^{p} e_i(z_1, ..., z_r)^{m_i}$, the product of elementary symmetric polynomials. Expanding this product into the sum of monomials we see that each monomial $z_1^{l_1} \cdots z_r^{l_r}$ occurs in $e_\nu(z_1, ..., z_r)$ with coefficient $M_{\lambda\nu}$, where λ is the partition $(l_1, ..., l_r)$. It remains to expand the RHS of (3.1) into the sum of polynomials e_ν, which is done in a straightforward way. ▮

The quantities $M_{\lambda\nu}$ play an important role in combinatorics, representation theory of symmetric groups, and classical theory of symmetric polynomials (see, e.g., [13]). The description of all pairs (λ, ν) such that $M_{\lambda\nu} > 0$ is known as the Gale–Ryser theorem. To state it we need two more notions about partitions, viz., conjugate partitions and dominance order.

Each partition ν can be written in a *normal form* $\nu = (n_1, ..., n_s)$, where $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_s > 0$. If ν is expressed in this way then its *diagram* is defined as the finite set

$$\{(k, j) \in \mathbf{Z}^2 : 1 \leqslant k \leqslant s,\ 1 \leqslant j \leqslant n_k\}.$$

The *conjugate* partition ν' is defined by the condition that its diagram be the transpose of the diagram of ν; in other words, if $\nu = (1^{m_1}, 2^{m_2}, ..., p^{m_p})$ then ν' is written in a normal form as $\nu' = (m_1 + m_2 + \cdots + m_p, m_2 + \cdots + m_p, ..., m_p)$.

The *dominance partial order* on partitions is defined as follows. Suppose partitions λ and ν are written as $\lambda = (l_1, ..., l_r)$ and $\nu = (n_1, ..., n_r)$ with $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_r$, $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r$. We say that λ *is dominated by* ν and write $\lambda \leqslant \nu$ if $l_1 + \cdots + l_k \leqslant n_1 + \cdots + n_k$ for $k = 1, ..., r-1$ and $l_1 + \cdots + l_r = n_1 + \cdots + n_r$.

PROPOSITION 3.4 (Gale–Ryser theorem [16]). *We have $M_{\lambda\nu} > 0$ if and only if $\lambda \leqslant \nu'$. Furthermore, $M_{\nu'\nu} = 1$.*

To illustrate the use of the Gale–Ryser theorem let us give another proof of Theorem 1.3. Let $\lambda = (l_1, ..., l_r)$ be a partition. By (3.6), we see that $N(l_1, ..., l_r) > 0$ if and only if $\lambda \leqslant \nu'$ for some partition $\nu = (1^{m_1}, 2^{m_2}, ..., p^{m_p})$ with $m_1 = 0$. By the Gale–Ryser theorem, Theorem 1.3 is an immediate consequence of the following combinatorial lemma.

LEMMA 3.5. *A partition $\lambda = (l_1, ..., l_r)$ is dominated by ν' for some $\nu = (1^{m_1}, 2^{m_2}, ...)$ with $m_1 = 0$ if and only if λ satisfies* (1.4).

Proof. Let ν' have a normal form $\nu' = (n'_1, n'_2, ..., n'_p)$. Then the condition that ν have no parts equal to 1 means that $n'_1 = n'_2$. We can assume that λ is also written in a normal form, i.e., $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_r$; then (1.4) means that $l_1 \leqslant l_2 + l_3 + \cdots + l_r$.

Now, if λ is dominated by ν' as above then we have

$$l_1 \leqslant n'_1 \leqslant n'_2 + \cdots + n'_p \leqslant l_2 + \cdots + l_r.$$

Conversely, if $l_1 \leqslant l_2 + l_3 + \cdots + l_r$ then we obviously have $\lambda \leqslant (l_1, l_1, l_2 + \cdots + l_r - l_1)$, and we are done. ∎

In many special cases the quantities $M_{\lambda\nu}$ can be evaluated explicitly, which leads to a more explicit formula for the degree $N(l_1, l_2, ..., l_r)$. Probably, the most important special case is the following. We say that the matrix format $(l_1+1)\times\cdots\times(l_r+1)$ is *boundary* if one of the inequalities (1.4) becomes an equality; without loss of generality we can assume that $l_1 = l_2 + \cdots + l_r$.

COROLLARY 3.6. *The degree of the hyperdeterminant of the boundary format is equal to*

$$N(l_2+\cdots+l_r, l_2, ..., l_r) = (l_2+\cdots+l_r+1)\binom{l_2+\cdots+l_t}{l_2, ..., l_r} = \frac{(l_1+1)!}{l_2!\cdots l_r!}. \tag{3.7}$$

Proof. We will use the following obvious combinatorial statement.

LEMMA 3.7. *If* $\lambda = (l_2, ..., l_r)$ *and* $\nu = (1^{l_2+\cdots+l_r})$ *then* $M_{\lambda\nu} = \binom{l_2+\cdots+l_r}{l_2, ..., l_r}$.

It is easy to see that in the boundary case there is exactly one summand in (3.6), and it corresponds to $\nu = (2^{l_2+\cdots+l_r})$. It remains to observe that

$$M_{\lambda\nu} = \binom{l_2+\cdots+l_r}{l_2, ..., l_r}.$$

But this follows at once from Lemma 3.7 because every (0, 1)-matrix contributing to $M_{\lambda\nu}$ has all the entries in the first row equal to 1. ∎

Note that for $r=2$ the boundary format is just that of ordinary square matrices, and (3.7) expresses the fact that the (ordinary) determinant of an $(l+1)\times(l+1)$-matrix has degree $l+1$.

It is natural to call the matrix format $(l_1+1)\times\cdots\times(l_r+1)$ with $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_r$ *subboundary* if $l_1 = l_2+\cdots+l_r-1$.

COROLLARY 3.8. *The degree of the hyperdeterminant of the subboundary format is equal to*

$$N(l_2+\cdots+l_r-1, l_2, ..., l_r) = 2\binom{l_2+\cdots+l_r}{l_2, ..., l_r} e_2(l_2, ..., l_r). \tag{3.8}$$

Here $e_2(l_2, ..., l_r) = \sum_{2\leqslant i<i'\leqslant r} l_i l_{i'}$ is the second elementary symmetric polynomial.

Proof. As for the boundary format it is easy to see that there is exactly one summand in (3.6) corresponding to $\nu = (2^{l_2+\cdots+l_r-2}, 3^1)$. It remains to show that

$$M_{\lambda\nu} = \frac{(l_2+\cdots+l_r-2)!}{l_2!\cdots l_r!} e_2(l_2, ..., l_r).$$

Clearly, every (0, 1)-matrix contributing to $M_{\lambda\nu}$ has all the entries in the first row equal to 1. Consider now the column with sum 3, i.e., containing three units. We know that one of these units lies in the first row. Decomposing the set of our (0, 1)-matrices into the subsets according to the location of the remaining two units and using Lemma 3.7 we see that

$$M_{\lambda\nu} = \sum_{2 \leqslant i < i' \leqslant r} \binom{l_2 + \cdots + l_r - 2}{l_2, ..., l_i - 1, ..., l_{i'} - 1, ..., l_r},$$

which readily implies our statement. ∎

COROLLARY 3.9. *The degree of the hyperdeterminant of the cubic matrix of format* $(l+1) \times (l+1) \times (l+1)$ *is equal to*

$$N(l, l, l) = \sum_{0 \leqslant k \leqslant l/2} \frac{(k+l+1)!}{k!^3 (l-2k)!} \cdot 2^{l-2k}. \tag{3.9}$$

Proof. We have $\lambda = (l, l, l)$. By the Gale–Ryser theorem, the partitions ν contributing to (3.6) have the form $\nu_k = (2^{3k}, 3^{l-2k})$ for $0 \leqslant k \leqslant l/2$. Clearly, $M_{\lambda\nu_k} = M_{((2k)^3)(2^{3k})}$ because every (0, 1)-matrix contributing to $M_{\lambda\nu_k}$ has all the unit entries in the first $(l-2k)$ columns. It is also easy to see that

$$M_{((2k)^3)(2^{3k})} = \binom{3k}{k, k, k} \tag{3.10}$$

because every $3 \times 3k$ matrix contributing to $M_{((2k)^3)(2^{3k})}$ is determined by a disjoint decomposition of the set of columns into three k-element subsets C_{12}, C_{13}, C_{23}, where C_{ij} is the set of columns with units in the ith and jth rows. Substituting (3.10) into (3.6) we get (3.9). ∎

For $l = 1, 2, 3$ the degree (3.9) is equal to 4, 36, and 272, respectively. It seems that the sum in (3.9) cannot be simplified.

Our last application of the formula (3.6) is the expression for the degree of the hyperdeterminant of the r-dimensional matrix of format $2 \times 2 \times \cdots \times 2$. Denote this degree by $N((1^r))$.

COROLLARY 3.10. *The exponential generating function for the sequence* $N((1^r))$ *is given by*

$$\sum_{r \geqslant 0} N((1^r))\, z^r/r! = e^{-2z}(1-z)^{-2}. \tag{3.11}$$

Proof. It follows from (3.6) and Lemma 3.7 that

$$N((1^r)) = r! \sum_{\nu} \frac{(m_2 + \cdots + m_r + 1)!}{\prod_{i \geqslant 2} [(i-2)!\, i]^{m_i}\, m_i!}, \tag{3.12}$$

the sum over all partitions $\nu=(2^{m_2}, 3^{m_3}, \ldots)$ with $2m_2+3m_3+\cdots=r$. Therefore, we have

$$\sum_{r\geqslant 0} N((1^r))\, z^r/r! = \sum \frac{(m_2+\cdots+m_p+1)!\, z^{2m_2+3m_3+\cdots}}{\prod_{i\geqslant 2}[(i-2)!i]^{m_i}\, m_i!},$$

the sum over all finite sequences $(m_2, \ldots, m_p)$ of nonnegative integers. The transformations similar to those used in the proof of Theorem 3.1 imply that the latter sum is equal to $(1-\sum_{i\geqslant 2}(z^i/(i-2)!i))^{-2}$. It remains to observe that

$$1-\sum_{i\geqslant 2}\frac{z^i}{(i-2)!i}=(1-z)\,e^z$$

since $1/(i-2)!i=1/(i-1)!-1/i!$. ∎

In particular, for $r=2, 3, 4, 5, 6$ the degree is respectively 2, 4, 24, 128, 880 and then grows very fast.

4. Hyperdeterminant of Boundary Format and Its Applications

Let $A=(a_{i_1\cdots i_r i_0})_{0\leqslant i_k\leqslant l_k}$ be an $(r+1)$-dimensional matrix of *boundary* format $(l_1+1)\times\cdots\times(l_r+1)\times(l_0+1)$, i.e., $l_0=l_1+l_2+\cdots+l_r$. In this section we show that $\mathrm{Det}(A)$ in this case can be interpreted as the *resultant* of the system of multilinear forms. This will lead us to an explicit formula (actually a number of explicit formulas) for $\mathrm{Det}(A)$ similar to the classical Sylvester formula for the resultant of two binary forms.

Let us introduce r groups of variables $x^{(k)}=(x_0^{(k)}, x_1^{(k)}, \ldots, x_{l_k}^{(k)})$ for $1\leqslant k\leqslant r$. Let $S(m_1, \ldots, m_r)$ denote the space of all polynomials in variables $x^{(1)}, \ldots, x^{(r)}$ which are homogeneous of degree m_k in variables of each group $x^{(k)}$. We will view our matrix A as a collection of (l_0+1) multilinear forms $f_0, f_1, \ldots, f_{l_0}\in S(1, 1, \ldots, 1)$ corresponding to the slices of A in the 0th direction:

$$f_{i_0}=\sum_{i_1,\ldots,i_r} a_{i_1\cdots i_r i_0} x_{i_1}^{(1)}\cdots x_{i_r}^{(r)}. \tag{4.1}$$

Theorem 4.1. *The hyperdeterminant* $\mathrm{Det}(A)$ *of the matrix of boundary format is equal to the resultant of the system of multilinear forms* f_0, $f_1, \ldots, f_{l_0}$, *i.e., A is degenerate if and only if the system of multilinear equations*

$$f_0(x)=f_1(x)=\cdots=f_{l_0}(x)=0 \tag{4.2}$$

has a non-trivial solution.

Proof. The "only if" part is obvious because the compatibility of (4.2) is one of the conditions defining degeneracy (see Section 1). The "if" part is clear from the fact that the compatibility of (4.2) is already a non-trivial condition on matrix entries, which is seen by dimension count.

In fact, Theorem 4.1 is a special case of a general fact on resultants proven in [10, Corollary 4.3, Example 4.5]. ∎

Note that in the case $l_1 = \cdots = l_r = 1$, $l_0 = r$ the hyperdeterminant is a special case of Dixon's resultant [5]. (Dixon considered systems of equations of arbitrary (multi)degrees on the product of projective lines.)

Analyzing the conditions of degeneracy (see Proposition 1.1) one can easily generalize Theorem 4.1 to the case when $l_0 > l_1 + l_2 + \cdots + l_r$. In this case we define the multilinear forms f_{i_0} by the same formula, (4.1).

THEOREM 4.1′. *Suppose that $l_0 \geqslant l_1 + l_2 + \cdots + l_r$. Then a matrix A of format $(l_1+1) \times \cdots \times (l_r+1) \times (l_0+1)$ is degenerate if and only if the system* (4.2) *has a non-trivial solution. The subvariety of degenerate matrices has codimension $l_0 - (l_1 + l_2 + \cdots + l_r) + 1$.*

Till the end of this section we assume that $l_0 = l_1 + l_2 + \cdots + l_r$. Let

$$m_k = l_1 + l_2 + \cdots + l_{k-1}, \qquad k = 1, ..., r \tag{4.3}$$

(with the convention $m_1 = 0$). We associate to our matrix A the linear operator

$$\partial_A \colon S(m_1, m_2, ..., m_r)^{l_0+1} \to S(1+m_1, 1+m_2, ..., 1+m_r)$$

given by $\partial_A(g_0, ..., g_{l_0}) = \sum_{i=0}^{l_0} f_i g_i$.

PROPOSITION 4.2. *Each of the spaces $S(m_1, m_2, ..., m_r)^{l_0+1}$ and $S(1+m_1, 1+m_2, ..., 1+m_r)$ has the dimension $N = (l_0+1)!/l_1!\, l_2! \cdots l_r!$.*

Proof. This follows at once from the standard fact that the number of monomials of degree m in $l+1$ variables, i.e., $\dim(S^m(\mathbf{C}^{l+1}))$ is equal to $\binom{l+m}{l}$. ∎

Let us choose in each of the spaces $S(m_1, m_2, ..., m_r)^{l_0+1}$ and $S(1+m_1, 1+m_2, ..., 1+m_r)$ the basis consisting of monomials. We will denote by the same symbol ∂_A the matrix of the operator ∂_A in these bases. By Proposition 4.2, this matrix is square.

THEOREM 4.3. *We have* $\mathrm{Det}(A) = \det(\partial_A)$.

Proof. First suppose that A is degenerate. By Theorem 4.1, this means that the system (4.2) has a non-trivial solution x. This obviously implies that each polynomial $h \in \mathrm{Im}(\partial_A)$ vanishes at x. Therefore, ∂_A is not onto, and so $\det(\partial_A) = 0$. This implies that the polynomial $\det(\partial_A)$ is divisible by $\mathrm{Det}(A)$.

Clearly, each entry of the matrix ∂_A is a linear form in matrix entries of A. Comparing Proposition 4.2 with Corollary 3.6 we see that the polynomials $\mathrm{Det}(A)$ and $\det(\partial_A)$ have the same degree. Therefore, to prove Theorem 4.3 it remains to establish the following lemma.

LEMMA 4.4. *The polynomial* $\det(\partial_A)$ *is non-zero, and it is irreducible over* $\mathbf{Z}$.

We will give two different proofs of Lemma 4.4 because we believe each of them is of independent interest.

First proof of Lemma 4.4. It suffices to exhibit a matrix E with integer entries such that $\det(\partial_E) = \pm 1$ (recall that $\mathrm{Det}(E)$ is defined only up to sign). Let E be the matrix whose entry $a_{i_1 \cdots i_r i_0}$ is equal to 1 if $i_0 = i_1 + \cdots + i_r$ and is equal to 0 otherwise. To show that $\det(\partial_E) = \pm 1$ it is enough to establish the following fact.

PROPOSITION 4.5. *The matrix* ∂_E *becomes triangular with units along the main diagonal under a suitable ordering of its rows and columns.*

Proof. First we give an explicit description of the matrix ∂_A. We identify the set of all monomials of degree m in $(l+1)$ variables $x_0, ..., x_l$ with the set of their exponent vectors

$$\Delta^l(m) = \left\{ b = (b_0, ..., b_l) \in \mathbf{Z}^{l+1} : b_i \geqslant 0, \sum b_i = m \right\}.$$

Thus the set of all monomials in $S(m_1, m_2, ..., m_r)$ is identified with the set

$$D = D(l_1, ..., l_r) = \Delta^{l_1}(m_1) \times \cdots \times \Delta^{l_r}(m_r),$$

and the set of all monomials in $S(1 + m_1, 1 + m_2, ..., 1 + m_r)$ with

$$R = R(l_1, ..., l_r) = \Delta^{l_1}(1 + m_1) \times \cdots \times \Delta^{l_r}(1 + m_r).$$

Now the rows of ∂_A are labeled by the set R, and the columns are labeled by $C = C(l_1, ..., l_r) = D \times [0, l_0]$, where $[0, l_0] = \{0, 1, ..., l_0\}$. We will denote a matrix entry of ∂_A by

$$\langle \mathbf{c}; i_0 \mid \mathbf{b} \rangle = \langle c^{(1)}, ..., c^{(r)}; i_0 \mid b^{(1)}, ..., b^{(r)} \rangle,$$

where $\mathbf{c} = (c^{(1)}, ..., c^{(r)}) \in D$, $i_0 \in [0, l_0]$, $\mathbf{b} = (b^{(1)}, ..., b^{(r)}) \in R$. We say that $\mathbf{b}$ *covers* $\mathbf{c}$ if $b^{(k)} - c^{(k)}$ has the form e_{i_k} for each $k = 1, ..., r$, where e_i is a vector with the ith component 1 and zeros elsewhere; in this case we write $\mathbf{b} \to \mathbf{c}$ or $\mathbf{b} \xrightarrow{i_1 \cdots i_r} \mathbf{c}$. By definition, $\langle \mathbf{c}; i_0 \mid \mathbf{b} \rangle = 0$ unless $\mathbf{b}$ covers $\mathbf{c}$, and if $\mathbf{b} \xrightarrow{i_1 \cdots i_r} \mathbf{c}$ then $\langle \mathbf{c}; i_0 \mid \mathbf{b} \rangle = a_{i_1 \cdots i_r i_0}$.

In particular, we see that ∂_E is a (0, 1)-matrix, and its entry $\langle \mathbf{c}; i_0 \mid \mathbf{b} \rangle$ is equal to 1 if and only if $\mathbf{b} \xrightarrow{i_1 \cdots i_r} \mathbf{c}$ for some $i_1, ..., i_r$ such that $i_1 + \cdots + i_r = i_0$. In this case we say that $\mathbf{b} \in R$ and $(\mathbf{c}; i_0) \in C$ are *incident* to each other.

For $0 \leqslant k \leqslant r$ we let

$$R_k = R_k(l_1, ..., l_r) = \{\mathbf{b} \in R : b_{l_p}^{(p)} > 0 \text{ for } k < p \leqslant r,\ b_{l_k}^{(k)} = 0\}$$

(for $k = 0$ the last condition is empty). For $1 \leqslant k \leqslant r$ we let

$$C_k = C_k(l_1, ..., l_r) = \{(\mathbf{c}; i_0) \in C : c_{l_p}^{(p)} > 0 \text{ for } k < p \leqslant r,\ c_{l_k}^{(k)} = 0,\ i_0 < l_0\};$$

also let

$$C_0 = C_0(l_1, ..., l_r) = \{(\mathbf{c}; i_0) \in C : i_0 = l_0,\ \mathbf{c} \text{ arbitrary}\}.$$

LEMMA 4.6. (a) *We have* $R = \bigcup_{0 \leqslant k \leqslant r} R_k$, $C = \bigcup_{0 \leqslant k \leqslant r} C_k$, *both unions disjoint.*

(b) *If* $\mathbf{b} \in R_k$ *is incident to* $(\mathbf{c}; i_0) \in C_p$ *then* $p \geqslant k$.

(c) *For every* $(\mathbf{c}; i_0) \in C_0$ *there is exactly one* $\mathbf{b} \in R$ *which is incident to* $(\mathbf{c}; i_0)$.

(d) *For every* $k = 1, ..., r$ *there are natural bijections*

$$R_k(l_1, ..., l_r) \to R(l_1, ..., l_{k-1}, l_k - 1, l_{k+1}, ..., l_r),$$
$$C_k(l_1, ..., l_r) \to C(l_1, ..., l_{k-1}, l_k - 1, l_{k+1}, ..., l_r)$$

such that elements $\mathbf{b} \in R_k(l_1, ..., l_r)$ *and* $(\mathbf{c}; i_0) \in C_k(l_1, ..., l_r)$ *are incident to each other if and only if their images are incident to each other.*

Proof. The statements (a) to (c) are immediate consequences of the definitions. The bijections in (d) are defined as follows. The image of $\mathbf{b} \in R_k(l_1, ..., l_r)$ is obtained from $\mathbf{b}$ by forgetting the coordinate $b_{l_k}^{(k)}$ and subtracting 1 from $b_{l_p}^{(p)}$ for $k < p \leqslant r$; the image of $(\mathbf{c}; i_0) \in C_k(l_1, ..., l_r)$ is defined in exactly the same way (with i_0 remaining unchanged). The statement (d) now also follows at once from the definitions. ∎

We can now easily complete the proof of Proposition 4.5. Choose an ordering of R so that elements of R_k will precede elements of R_p for $k < p$ (and similarly for C). By Lemma 4.6(b), under such orderings the matrix ∂_E becomes block triangular with $(r+1)$ diagonal blocks, the kth block being the incidence matrix for the incidence relation between R_k and C_k, where $0 \leqslant k \leqslant r$. By Lemma 4.6(c), the 0th block becomes the identity matrix under a suitable ordering of R_0 and C_0. But if $1 \leqslant k \leqslant r$ then by Lemma 4.6(d), the kth block of ∂_E coincides with the matrix of the same type ∂_E corresponding to E of the format $(l_1+1) \times \cdots \times (l_{k-1}+1) \times l_k \times (l_{k+1}+1) \times \cdots \times (l_r+1) \times l_0$. Using induction on $l_0 = l_1 + \cdots + l_r$ we can assume that all the diagonal blocks of ∂_E can be made unitriangular by a permutation of rows and columns. Therefore, the same is true for ∂_E itself. This proves Proposition 4.5, and hence Lemma 4.4 and Theorem 4.3. ▌

The matrix E can be viewed as a multidimensional analog of the identity matrix. In fact, one can show easily that the corresponding system (4.2) has only a trivial solution. To see this we represent each vector $x^{(k)} = (x_0^{(k)}, x_1^{(k)}, ..., x_{l_k}^{(k)})$ by a "generating" polynomial $P^{(k)}(t) = \sum_{i=0}^{l_k} x_i^{(k)} t^i$. Then the system (4.2) for $A = E$ can be written as

$$P^{(1)}(t)\, P^{(2)}(t) \cdots P^{(r)}(t) = 0,$$

which implies that some $P^{(k)}(t)$ is zero polynomial, i.e., $x^{(k)} = 0$. It would be interesting to find an analogous matrix in the "interior" case when $l_0 < l_1 + \cdots + l_r$.

Proposition 4.5 has an amazing combinatorial corollary.

COROLLARY 4.7. *There exists exactly one bijection $\phi: R \to C$ such that $\phi(\mathbf{b})$ is incident to $\mathbf{b}$ for each $\mathbf{b} \in R$.*

The bijection ϕ from Corollary 4.7 and its inverse can be explicitly constructed as follows. For $\mathbf{b} = (b^{(1)}, ..., b^{(r)}) \in R$ we define the indices $i_1, ..., i_r$ successively: if $i_1, ..., i_{k-1}$ are already constructed, we define i_k as the minimal index such that

$$b_0^{(k)} + b_1^{(k)} + \cdots + b_{i_k}^{(k)} > i_1 + i_2 + \cdots + i_{k-1}$$

(the index i_1 is defined by $b^{(1)} = e_{i_1}$). We then define

$$\phi(\mathbf{b}) = (b^{(1)} - e_{i_1}, ..., b^{(r)} - e_{i_r};\; i_1 + \cdots + i_r). \tag{4.4}$$

It is easy to see that ϕ is a well-defined mapping from R to C such that $\phi(\mathbf{b})$ is incident to $\mathbf{b}$ for each $\mathbf{b} \in R$.

To show that ϕ is a desired bijection we construct the inverse mapping $\psi: C \to R$. For $(c^{(1)}, ..., c^{(r)}; i_0) \in C$ we define the indices $i_r, i_{r-1}, ..., i_1$ successively: if $i_r, ..., i_{k+1}$ are already constructed, we define i_k as the minimal index such that

$$(1 + c_0^{(k)}) + (1 + c_1^{(k)}) + \cdots + (1 + c_{i_k}^{(k)}) > i_0 - i_r - \cdots - i_{k+1}.$$

We then define

$$\psi(\mathbf{c}; i_0) = (c^{(1)} + e_{i_1}, ..., c^{(r)} + e_{i_r}). \tag{4.5}$$

It is straightforward to verify that ψ is well-defined, and both compositions $\psi \circ \phi$ and $\phi \circ \psi$ are identity mappings.

Second proof of Lemma 4.4. For $m_1, ..., m_r \in \mathbf{Z}$ let $\mathcal{O}(m_1, ..., m_r)$ be the sheaf on $P^{l_1} \times \cdots \times P^{l_r}$ introduced in Section 2 above. Clearly, the space of global sections $H^0(P^{l_1} \times \cdots \times P^{l_r}, \mathcal{O}(m_1, ..., m_r))$ is non-zero if and only if all $m_k \geqslant 0$, and in this case we have

$$H^0(P^{l_1} \times \cdots \times P^{l_r}, \mathcal{O}(m_1, ..., m_r)) = S(m_1, ..., m_r).$$

Now let $m_1, ..., m_r \in \mathbf{Z}$ and $f_0, ..., f_{l_0} \in S(1, 1, ..., 1)$ have the same meaning as above, i.e., they are given by (4.3) and (4.1), respectively. We associate to them the complex of sheaves $\mathcal{K}^\cdot = \mathcal{K}^\cdot(m_1, ..., m_r; A)$: $\{\mathcal{K}^{-l_0-1} \xrightarrow{\partial_A} \mathcal{K}^{-l_0} \xrightarrow{\partial_A} \cdots \xrightarrow{\partial_A} \mathcal{K}^{-1} \xrightarrow{\partial_A} \mathcal{K}^0\}$ on $P^{l_1} \times \cdots \times P^{l_r}$, where

$$\mathcal{K}^{-p} = \mathcal{O}(1 + m_1 - p, 1 + m_2 - p, ..., 1 + m_r - p) \otimes \Lambda^p(\mathbf{C}^{l_0+1}).$$

The differential $\partial_A: \mathcal{K}^{-p} \to \mathcal{K}^{-p+1}$ is defined as $\partial_A = \sum_{i=0}^{l_0} f_i \otimes \partial_i$, where each f_i is thought of as a multiplication operator, and ∂_i is a (super)derivation of the exterior algebra $\Lambda^\cdot(\mathbf{C}^{l_0+1})$ w.r.t. the ith standard basis vector.

Now suppose that A is nondegenerate. By Theorem 4.1, this means that $f_0, ..., f_{l_0}$ do not have a non-trivial common root. Clearly, this implies that the complex of sheaves $\mathcal{K}^\cdot(m_1, ..., m_r; A)$ is exact. On the other hand, consider the complex of vector spaces $K^\cdot = K^\cdot(m_1, ..., m_r; A)$ obtained from $\mathcal{K}^\cdot(m_1, ..., m_r; A)$ by taking global sections, i.e., $K^{-p} = S(1 + m_1 - p, 1 + m_2 - p, ..., 1 + m_r - p) \otimes \Lambda^p(\mathbf{C}^{l_0+1})$. Since $m_1 = 0$ it follows that $K^\cdot$ has only two non-zero terms K^0 and K^{-1}, and the corresponding map $\partial_A: K^{-1} \to K^0$ is identified with the linear operator $\partial_A: S(m_1, m_2, ..., m_r)^{l_0+1} \to S(1 + m_1, 1 + m_2, ..., 1 + m_r)$ from Theorem 4.3. Thus the statement of Lemma 4.4 that $\det(\partial_A) \neq 0$ means that $K^\cdot$ is exact, i.e., that the functor of taking global sections preserves the exactness of $\mathcal{K}^\cdot$. By a general criterion which we already used in Section 2, this is a consequence of the fact that all the terms of $\mathcal{K}^\cdot$ do not have higher cohomology. Therefore, to show that $\det(\partial_A) \neq 0$ it suffices to prove the following lemma.

LEMMA 4.8. *For each* $p = 0, 1, ..., l_0 + 1$ *and* $j > 0$ *we have* $H^j(P^{l_1} \times \cdots \times P^{l_r}, \mathcal{O}(1 + m_1 - p, 1 + m_2 - p, ..., 1 + m_r - p)) = 0$.

Proof. By the Künneth formula, we have

$$H^j(P^{l_1} \times \cdots \times P^{l_r}, \mathcal{O}(1+m_1-p, 1+m_2-p, ..., 1+m_r-p))$$
$$= \bigoplus_{j_1+\cdots+j_r=j} \bigotimes_{k=1}^{r} H^{j_k}(P^{l_k}, \mathcal{O}(1+m_k-p)).$$

It is known that the cohomology $H^j(P^l, \mathcal{O}(m))$ is non-zero only in the following cases: $m \geqslant 0$, $j=0$ or $m < -l$, $j=l$ (see [14, Chap. 1, Sect. 1]). In particular, if $-l \leqslant m < 0$ then $H^j(P^l, \mathcal{O}(m)) = 0$ for all j.

Now for $p=0$ or $p=1$ we have $1+m_k-p \geqslant 0$ for all k, hence $H^{k_k}(P^{l_k}, \mathcal{O}(1+m_k-p)) = 0$ for $j_k > 0$, which implies our statement. So we can assume that $1 < p \leqslant l_0 + 1$. Then we have

$$1 + l_1 + \cdots + l_{k-1} < p \leqslant 1 + l_1 + \cdots + l_{k-1} + l_k \tag{4.6}$$

for some $k = 1, ..., r$. Using (4.3) we can rewrite (4.6) as

$$-l_k \leqslant 1 + m_k - p < 0.$$

This implies that $H^{j_k}(P^{l_k}, \mathcal{O}(1+m_k-p)) = 0$ for all j_k, which proves our lemma. ∎

To complete the proof of Lemma 4.4 it remains to show that the polynomial $\det(\partial_A)$ is irreducible over $\mathbf{Z}$ (under the choice of bases described above). This is proven by the same arguments as those in [6, Sect. 2E]. Lemma 4.4, and hence Theorem 4.3, is proven. ∎

Note that in the case $l_1 = \cdots = l_r = 1$ Theorem 4.3 is a special case of the formula due to Dixon [5].

EXAMPLE 4.9. Let $r=3$ and $l_0 = 2$, $l_1 = l_2 = 1$, i.e., $A = (a_{ijk})$ $(0 \leqslant i, j \leqslant 1, 0 \leqslant k \leqslant 2)$ is a 3-dimensional matrix of format $2 \times 2 \times 3$. Let A_0 and A_1 be two slices of A in the second direction, i.e.,

$$A_j = \begin{bmatrix} a_{0j0} & a_{0j1} & a_{0j2} \\ a_{1j0} & a_{1j1} & a_{1j2} \end{bmatrix} \qquad (j = 0, 1).$$

Then the matrix ∂_A can be written as the block 6×6 matrix

$$\partial_A = \begin{bmatrix} A_0 & 0 \\ A_1 & A_0 \\ 0 & A_1 \end{bmatrix}. \tag{4.7}$$

(This corresponds to the following ordering of the sets R and C,

$$R = \{(0,0), (1,0), (0,1), (1,1), (0,2), (1,2)\},$$
$$C = \{(0;0), (0;1), (0;2), (1;0), (1;1), (1;2)\},$$

where (a, b) stands for the vector $((1-a, a), (2-b, b)) \in \Delta^1(1) \times \Delta^1(2) = R$, and $(c; k)$ stands for $((1-c, c); k) \in \Delta^1(1) \times [0, 2] = C$.) By Theorem 4.3, we have $\mathrm{Det}(A) = \det(\partial_A)$. This polynomial can be rewritten in many different ways. For instance, taking the Laplace expansion of $\det(\partial_A)$ in the three first columns we see that

$$\mathrm{Det}(A) = \det \begin{bmatrix} a_{000} & a_{001} & a_{002} \\ a_{100} & a_{101} & a_{102} \\ a_{010} & a_{011} & a_{012} \end{bmatrix} \det \begin{bmatrix} a_{100} & a_{101} & a_{102} \\ a_{010} & a_{011} & a_{012} \\ a_{110} & a_{111} & a_{112} \end{bmatrix}$$

$$- \det \begin{bmatrix} a_{000} & a_{001} & a_{002} \\ a_{100} & a_{101} & a_{102} \\ a_{110} & a_{111} & a_{112} \end{bmatrix} \det \begin{bmatrix} a_{000} & a_{001} & a_{002} \\ a_{010} & a_{011} & a_{012} \\ a_{110} & a_{111} & a_{112} \end{bmatrix}. \tag{4.8}$$

EXAMPLE 4.10. Consider a generalization of the previous example: $r = 3$, $l_0 = l$, $l_1 = l - 1$, $l_2 = 1$ for some $l \geqslant 1$, so A is a 3-dimensional matrix of format $l \times 2 \times (l+1)$. As before, let A_0 and A_1 be two slices of A in the second direction, i.e., A_0 and A_1 are $l \times (l+1)$ matrices. It is easy to see that ∂_A is a square matrix of order $l(l+1)$ which can be written in the form generalizing (4.7), i.e., as a block matrix with blocks of size $l \times (l+1)$, where for $c = 1, ..., l$ the cth column is $[0, ..., 0, A_0, A_1, 0, ..., 0]^t$ with A_0 at the cth place.

In fact, an expression of type (4.8) can be given for an arbitrary boundary format. To do this we introduce some terminology. For every subset $\Delta \subset [0, l_1] \times \cdots \times [0, l_r]$ of cardinality $l_0 + 1$ we define the polynomial $[\Delta] = [\Delta](A)$ to be

$$[\Delta] = \det(a_{\mathbf{i} i_0}) \qquad (\mathbf{i} \in \Delta,\ 0 \leqslant i_0 \leqslant l_0) \tag{4.9}$$

(note that if we do not specify an ordering of Δ then $[\Delta]$ is defined only up to sign).

Let R and D have the same meanings as above. We say that a mapping $\pi: R \to D$ is a *covering* if it satisfies two conditions:

(1) An element $\mathbf{b}$ covers $\pi(\mathbf{b})$ for each $\mathbf{b} \in R$.

(2) For each $\mathbf{c} \in D$ the subset $\pi^{-1}(\mathbf{c}) \subset R$ has cardinality $l_0 + 1$.

To every covering $\pi: R \to D$ and every $\mathbf{c} = (c^{(1)}, ..., c^{(r)}) \in D$ we associate the subset $\Delta_\pi(\mathbf{c}) \subset [0, l_1] \times \cdots \times [0, l_r]$ of cardinality $l_0 + 1$ in the following way:

$$\Delta_\pi(\mathbf{c}) = \{(i_1, ..., i_r) : (c^{(1)} + e_{i_1}, ..., c^{(r)} + e_{i_r}) \in \pi^{-1}(\mathbf{c})\}.$$

PROPOSITION 4.11. *The hyperdeterminant* Det(A) *can be written as* $\sum_\pi \pm \prod_{\mathbf{c} \in D} [\Delta_\pi(\mathbf{c})]$, *the sum over all coverings* $\pi: R \to D$.

Proof. Consider the Laplace expansion of $\det(\partial_A)$ corresponding to the following grouping of columns of ∂_A: we join together the columns $(\mathbf{c}; i_0)$ having the same component $\mathbf{c}$. By definition, summands in the Laplace expansion correspond to coverings $\pi: R \to D$, the summand corresponding to a covering π being just $\pm \prod_{\mathbf{c} \in D} [\Delta_\pi(\mathbf{c})]$. ∎

Note that the signs in Proposition 4.11 are calculated directly once we specify orderings of all subsets Δ; since Det(A) itself is defined only up to sign, we can actually compute only the ratio of signs for every two summands.

Remarks 4.12. (a) Even the existence of coverings $\pi: R \to D$ is a nontrivial combinatorial fact. Obviously, this implies (and in fact is equivalent to) the following property: for every subset $\Xi \subset D$ we have

$$\#\{\mathbf{b} \in R : \mathbf{b} \text{ covers } \mathbf{c} \text{ for some } \mathbf{c} \in \Xi\} \geqslant (l_0 + 1) \cdot \#\Xi.$$

(b) Proposition 4.11 has the following geometric interpretation. Each multilinear form f_i $(i = 0, ..., l_0)$ can be thought of as a linear form on the space $V = \mathbf{C}^{l_1+1} \otimes \cdots \otimes \mathbf{C}^{l_r+1}$, and so a generic matrix A defines a vector subspace $\xi_A = \{v \in V : f_0(v) = \cdots = f_{l_0}(v) = 0\} \subset V$ of codimension $l_0 + 1$. The polynomials $[\Delta]$ are *dual Plücker coordinates* of ξ_A (cf. [10, Sect. 2A). Therefore, Proposition 4.11 expresses Det(A) as a homogeneous function on the Grassmann variety of subspaces of codimension $l_0 + 1$ in V. By definition, this function is the *Chow form* of the subvariety $P^{l_1} \times \cdots \times P^{l_r} \subset P(V)$ (cf. [10, Example 4.5]).

(c) By [10, Theorem 5.3], the extreme terms in the expansion of Proposition 4.11 correspond to regular triangulations of the convex polytope $\Delta^{l_1} \times \cdots \times \Delta^{l_r}$, the product of several simplices. This correspondence reveals some interesting combinatorial and geometric properties of triangulations. They will be studied in a forthcoming publication.

Now we give another geometric interpretation of Det(A). Let V' be the space of all matrices of format $(l_2 + 1) \times \cdots \times (l_r + 1) \times (l_0 + 1)$. Let ∇' be the variety of all degenerate matrices in V', and let $X' \subset P(V')$ be the projectivization of ∇'. By Theorem 4.1′, X' has codimension $l_0 - (l_2 + \cdots + l_r) + 1 = l_1 + 1$ in $P(V')$. Consider its Chow form $R_{X'}$: by definition, it is a homogeneous regular function on the (cone over) Grassmann variety $G(l_1+1, V')$ which defines the hypersurface formed by subspaces $\xi \in G(l_1+1, V')$ which meet X' (see [10, Sect. 2). We will represent $R_{X'}$ as a polynomial in Plücker coordinates $[\Omega]$ on the Grassmannian $G(l_1+1, V')$, where Ω runs the subsets of $[0, l_2] \times \cdots \times [0, l_r] \times [0, l_0]$ of cardinality (l_1+1). We will show that this Chow form is given by the hyperdeterminant of the boundary format $(l_1+1) \times \cdots \times (l_r+1) \times (l_0+1)$ (cf. Remark 4.12(b)).

To be more precise we represent a matrix A of this boundary format as the ordinary (2-dimensional) matrix $(a_{i_1\mathbf{j}})$ of format $(l_1+1) \times [(l_2+1) \cdots (l_r+1)(l_0+1)]$, where $i_1 \in [0, l_1]$, $\mathbf{j} = (i_2, ..., i_r, i_0) \in [0, l_2] \times \cdots \times [0, l_r] \times [0, l_0]$. Thus, A represents a linear operator $\tilde{A}: \mathbf{C}^{l_1+1} \to V'$. Clearly, for all A except the subvariety of codimension more than 1, we have $\mathrm{rk}(\tilde{A}) = l_1+1$, i.e., $\mathrm{Im}(\tilde{A}) \in G(l_1+1, V')$.

THEOREM 4.13. *The Chow form $R_{X'}$ evaluated at* $\mathrm{Im}(\tilde{A})$ *is equal to* $\mathrm{Det}(A)$.

Proof. Remembering all the definitions, we have only to prove that A is degenerate if and only if $\tilde{A}(x) \in \nabla'$ for some non-zero $x \in \mathbf{C}^{l_1+1}$. But this follows at once from characterizations of degenerate matrices given by Theorems 4.1 and 4.1′ (we apply Theorem 4.1 to A and Theorem 4.1′ to $\tilde{A}(x) \in V'$). ∎

Note that the format $(l_2+1) \times \cdots \times (l_r+1) \times (l_0+1)$ is (up to permutation of its r directions) an arbitrary format not satisfying (1.4), i.e., such that the variety ∇' of all degenerate matrices of this format is of codimension greater than 1. Using general properties of the Chow form [19, Sect. 36] we can extract from Theorem 4.13 the following description of ∇'.

COROLLARY 4.14. *A matrix A' of format $(l_2+1) \times \cdots \times (l_r+1) \times (l_0+1)$ is degenerate if and only if the hyperdeterminant* $\mathrm{Det}(A)$ *of boundary format $(l_1+1) \times (l_2+1) \times \cdots \times (l_0+1)$ vanishes whenever A has A' as a slice in the first direction.*

To make Theorem 4.13 more explicit, for every subset $\Omega \subset [0, l_2] \times \cdots \times [0, l_r] \times [0, l_0]$ of cardinality l_1+1 let us denote (with some abuse of notation) by $[\Omega] = [\Omega](A)$ the corresponding Plücker coordinate of $\mathrm{Im}(\tilde{A})$, i.e., the minor

$$[\Omega] = [\Omega](A) = \det(a_{i_1\mathbf{j}}), \qquad i_1 \in [0, l_1],\ \mathbf{j} \in \Omega \tag{4.10}$$

(like the polynomials $[\Delta]$ above, $[\Omega]$ is defined only up to sign). Then Theorem 4.13 means, in particular, that $\mathrm{Det}(A)$ can be expressed as a polynomial in these minors. Such an expression can be given quite explicitly in full analogy with Proposition 4.11. For this we need some more terminology.

Let $R' = \Delta^{l_2}(1+m_2) \times \cdots \times \Delta^{l_r}(1+m_r)$, where $m_2, ..., m_r$ are given by (4.3) (thus the set R of row indices for the matrix ∂_A is equal to $\Delta^{l_1}(1) \times R'$). Let $C = D \times [0, l_0]$ have the same meaning as above. We say that an element $(\mathbf{c}; i_0) = (c^{(1)}, c^{(2)}, ..., c^{(r)}; i_0) \in C$ *covers* an element $\mathbf{b}' = (b^{(2)}, ..., b^{(r)}) \in R'$ if $b^{(k)} - c^{(k)}$ has the form e_{i_k} for $k = 2, 3, ..., r$. We say that a mapping $\tau: C \to R'$ is a *covering* if it satisfies two conditions:

(1) An element $(\mathbf{c}; i_0)$ covers $\tau(\mathbf{c}; i_0)$ for each $(\mathbf{c}; i_0) \in C$.

(2) For each $\mathbf{b}' \in R'$ the subset $\tau^{-1}(\mathbf{b}') \subset C$ has cardinality $l_1 + 1$.

To every covering $\tau: C \to R'$ and every $\mathbf{b}' = (b^{(2)}, ..., b^{(r)}) \in R'$ we associate the subset $\Omega_\tau(\mathbf{b}') \subset [0, l_2] \times \cdots \times [0, l_r] \times [0, l_0]$ of cardinality $l_1 + 1$ in the following way:

$$\Omega_\tau(\mathbf{b}') = \{(i_2, ..., i_r, i_0) : (0, b^{(2)} - e_{i_2}, ..., b^{(r)} - e_{i_r}; i_0) \in \tau^{-1}(\mathbf{b}')\}.$$

PROPOSITION 4.15. *The hyperdeterminant* $\mathrm{Det}(A)$ *can be written as* $\sum \pm \prod_{\mathbf{b}' \in R'} [\Omega_\tau(\mathbf{b}')]$, *the sum over all coverings* $\tau: C \to R'$.

Proof. Consider the Laplace expansion of $\det(\partial_A)$ corresponding to the following grouping of rows of ∂_A: we join together the rows $(b^{(1)}, \mathbf{b}')$ having the same component $\mathbf{b}'$. By definition, summands in the Laplace expansion correspond to coverings $\tau: C \to R'$, the summand corresponding to a covering τ being just $\pm \prod_{\mathbf{b}' \in R'} [\Omega_\tau(\mathbf{b}')]$. ∎

Remark 4.16. Combining Theorem 4.3 with Proposition 1.10, we obtain $r!$ different determinantal formulas for the hyperdeterminant $\mathrm{Det}(A)$ of an $(r+1)$-dimensional boundary format. Namely, for each permutation σ of indices $1, ..., r, r+1$ leaving $(r+1)$ invariant we have

$$\mathrm{Det}(A) = \det(\partial_{\sigma(A)}), \tag{4.11}$$

where $\partial_{\sigma(A)}$ is a square matrix constructed as above (see the first proof of Lemma 4.4) but with respect to the "transpose" matrix $\sigma(A)$ instead of A. All the matrices $\partial_{\sigma(A)}$ are of the same order and consist of zeros and matrix entries of A. But their block structures can differ substantially, so it is far from obvious that they have the same determinant. In particular, the expressions for $\mathrm{Det}(A)$ as Chow forms given by Propositions 4.11 and 4.15 depend upon the choice of σ. Of course, two different expressions for the same Chow form can be transformed one to another by means of Plücker relations on the corresponding Grassmannian.

5. Schläfli's Method

In this section we study the method of computing hyperdeterminants due to L. Schläfli [18]. Although it does not provide an answer in general, it works in some important special cases and provides interesting additional information.

Let $A=(a_{i_0 i_1 \cdots i_r})$ be an $(r+1)$-dimensional matrix of format $(l_0+1)\times(l_1+1)\times\cdots\times(l_r+1)$. We associate to A the family of r-dimensional matrices $\tilde{A}(x)$ linearly depending on the auxiliary variables $x=(x_0, ..., x_{l_0})$:

$$\tilde{A}(x)_{i_1 \cdots i_r}=\sum_{i_0=0}^{l_0} a_{i_0 i_1 \cdots i_r} x_{i_0}. \tag{5.1}$$

In other words, $\tilde{A}$ is the linear operator

$$\mathbf{C}^{l_0+1}\to\mathbf{C}^{l_1+1}\otimes\cdots\otimes\mathbf{C}^{l_r+1}$$

naturally associated to A.

Let us assume that the numbers $l_1, ..., l_r$ satisfy (1.4), i.e., the r-dimensional hyperdeterminant of format $(l_1+1)\times\cdots\times(l_r+1)$ is defined. We associate to A the polynomial function $F_A(x)=\mathrm{Det}(\tilde{A}(x))$. This is a homogeneous form in $x_0, ..., x_{l_0}$ of degree $N(l_1, ..., l_r)$ (see Section 3). Denote by $\Delta(F_A)$ the discriminant of F_A. Recall that the discriminant $\Delta(\Phi)$ of a homogeneous polynomial $\Phi(x_1, ..., x_n)$ (form) of degree d is the irreducible polynomial in coefficients of Φ which vanishes if and only if the projective hypersurface $\{\Phi=0\}$ is singular. For binary forms of degree d it is given by the classical Sylvester formula (see, e.g., [19, Sect. 15]) and has degree (in coefficients of Φ) equal to $2d-2$. In the general case the degree of Δ equals $n(d-1)^{n-1}$; see, e.g., [18].

We consider $\Delta(F_A)$ as a polynomial in matrix entries of A. Therefore, its degree is equal to

$$\deg(\Delta(F_A))=(l_0+1)(N(l_1, ..., l_r)-1)^{l_0} N(l_1, ..., l_r). \tag{5.2}$$

Theorem 5.1. *The polynomial $\Delta(F_A)$ is divisible by the $(r+1)$-dimensional hyperdeterminant of the matrix A.*

Note that $\Delta(F_A)$ might be identically zero.

Proof. Suppose that $A=(a_{i_0 i_1 \cdots i_r})$ is degenerate; we have to show that the corresponding form $F_A(x)=\mathrm{Det}(\tilde{A}(x))$ has zero discriminant. Choose a point $(x^{(0)}, x^{(1)}, ..., x^{(r)})\in\mathscr{K}(A)$ (see Proposition 1.1). It is enough to show that F_A vanishes at $x^{(0)}$ with all its first derivatives. Denote

$$A_0=\tilde{A}(x^{(0)}),\ b_{i_1 \cdots i_r}=\left.\frac{\partial\,\mathrm{Det}(A)}{\partial a_{i_1 \cdots i_r}}\right|_{A=A_0}.$$

Then we have $F_A(x^{(0)}) = \mathrm{Det}(A_0)$, and

$$\left.\frac{\partial F_A(x)}{\partial x_{i_0}}\right|_{x=x^{(0)}} = \sum_{i_1, \ldots, i_r} a_{i_0 i_1 \cdots i_r} b_{i_1 \cdots i_r} \tag{5.3}$$

for all $i_0 = 0, 1, \ldots, l_0$. Clearly, $(x^{(1)}, \ldots, x^{(r)}) \in \mathscr{K}(A_0)$, hence $F_A(x^{(0)}) = 0$. If $b_{i_1 \cdots i_r} = 0$ for all $i_1, \ldots, i_r$ then by (5.3), $\partial F_A(x)/\partial x_{i_0}\,|_{x=x^{(0)}} = 0$ for all i_0, and we are done. So we can assume that some $b_{i_1 \cdots i_r}$ is non-zero. But this means that A_0 is a smooth point of the variety of degenerate matrices, and we can apply Proposition 1.2. By this proposition, we can assume that $b_{i_1 \cdots i_r} = x_{i_1}^{(1)} \cdots x_{i_r}^{(r)}$ for all $i_1, \ldots, i_r$. Substituting this into (5.3) and remembering the definition of $\mathscr{K}(A)$ we see that all the first partial derivatives of F_A vanish at $x^{(0)}$, which proves our theorem. ∎

Denote by $\nabla = \nabla(l_1, \ldots, l_r)$ the variety of all degenerate matrices of format $(l_1+1) \times \cdots \times (l_r+1)$; by definition, the projectivization of ∇ is the projective dual variety $X^\vee$ of $X = P^{l_1} \times \cdots \times P^{l_r}$. Let ∇_{sing} be the variety of singular points of ∇, and $X^\vee_{\mathrm{sing}}$ be the projectivization of ∇_{sing}. Let $c = c(l_1, \ldots, l_r)$ denote the codimension (i.e., the minimum of codimensions of irreducible components) of $X^\vee_{\mathrm{sing}}$ in the projective space $P(\mathbf{C}^{(l_1+1) \cdots (l_r+1)})$.

Analyzing the proof of Theorem 5.1 we get the following refinement.

THEOREM 5.2. *The ratio $G(A) = \Delta(F_A)/\mathrm{Det}(A)$ has the following form:*

(a) *If $l_0 + 1 < c(l_1, \ldots, l_r)$ then G is a non-zero constant.*

(b) *If $l_0 + 1 = c(l_1, \ldots, l_r)$ then $G(A) = \prod R_Z^{m_Z}(\mathrm{Im}(\tilde{A}))$, where Z ranges over irreducible components of $X^\vee_{\mathrm{sing}}$ having codimension $c(l_1, \ldots, l_r)$, R_Z is the Chow form of Z, and m_Z are some multiplicities.*

(c) *If $l_0 + 1 > c(l_1, \ldots, l_r)$ then G (and hence $\Delta(F_A)$) is identically zero.*

Here the Chow form R_Z in (b) is understood in the same sense as in Section 4: we have $R_Z(\mathrm{Im}(\tilde{A})) = 0$ if and only if the projectivization of $\tilde{A}(x^{(0)})$ lies in Z for some nonzero $x^{(0)} \in \mathbf{C}^{l_0+1}$.

Proof. First we establish the following.

LEMMA 5.3. *If $\Delta(F_A)$ is not identically zero then it is not divisible by* $\mathrm{Det}(A)^2$.

Proof. By definition, $X^\vee$ is the union of projective spaces P_x, $x \in X$, where P_x is the space of hyperplanes tangent to X at x. The codimension of P_x is equal to $\dim(X) + 1 = l_1 + \cdots + l_r + 1$. The vanishing of $\Delta(F_A)$ means that the image $\mathrm{Im}(\tilde{A}) = \tilde{A}(\mathbf{C}^{l_0+1})$ is tangent to ∇ at some nonzero point. Suppose that $\Delta(F_A)$ is divisible by $\mathrm{Det}(A)^2$. Then for any one-parameter algebraic family of $(r+1)$-dimensional matrices A_t such that $\operatorname{Im}(\tilde{A}_0)$ is tangent to ∇, the function $\Delta(F_{A_t})$ is divisible by t^2. We will show that this is impossible by constructing a suitable "generic" family.

Let B be a generic point of ∇ and $\xi \in X^\vee$ be the projectivization of B. We can assume that ξ lies on exactly one P_x (if this were generically not so then $X^\vee$ would not be a hypersurface). Consider the variety Z of all l_0-dimensional projective subspaces in $P(\mathbf{C}^{(l_1+1)\cdots(l_r+1)})$ which are tangent to $X^\vee$ at ξ. Since $l_0 \leqslant l_1 + \cdots + l_r$ it follows that a dense open part of Z is formed by subspaces which meet P_x only at ξ. Let L be a generic element of Z. Then L has a simple tangency with $X^\vee$. Now take a generic one-parameter family of matrices A_t such that L is the projectivization of $\operatorname{Im}(\tilde{A}_0)$. The simple tangency condition implies that the function $t \mapsto \Delta(F_{A_t})$ has a simple zero at $t=0$. This completes the proof of Lemma 5.3. ∎

Now we can easily complete the proof of Theorem 5.2. In the course of the proof of Theorem 5.1 we have actually shown that $\Delta(F_A)=0$ if and only if either $\operatorname{Det}(A)=0$ or $\tilde{A}(x^{(0)}) \in \nabla_{\text{sing}}$ for some non-zero $x^{(0)} \in \mathbf{C}^{l_0+1}$. Denote by W the variety of all matrices A such that $\operatorname{Im}(\tilde{A})$ meets ∇_{sing} at some non-zero point. Taking into account Lemma 5.3 we see that the ratio $G(A)$ may vanish only when $A \in W$.

Clearly, $\operatorname{codim}(W) > 1$ for $l_0 + 1 < c(l_1, ..., l_r)$, $\operatorname{codim}(W) = 1$ for $l_0 + 1 = c(l_1, ..., l_r)$, and W coincides with the whole matrix space for $l_0 + 1 > c(l_1, ..., l_r)$. Now all the assertions of our theorem follow at once from the definition of the Chow form. ∎

EXAMPLE 5.4. Let $r=2$, and ∇ be the space of degenerate $m \times m$ matrices, $m \geqslant 2$. The variety ∇_{sing} consists of matrices of rank $\leqslant (m-2)$ and has codimension 4. Therefore, the hyperdeterminant of a 3-dimensional matrix A of format $2 \times m \times m$ or $3 \times m \times m$ is equal to the discriminant of the binary (resp. ternary) form $\det \tilde{A}(x)$. Note that for matrices of format $2 \times 2 \times 2$ or $3 \times 3 \times 3$ we obtain in this way three different formulas for the hyperdeterminant corresponding to three different choices of a distinguished direction. For the format $2 \times 2 \times 2$ the hyperdeterminant is given by the formula (1.5). For each of the formats $2 \times m \times m$ and $3 \times m \times m$ we obtain from (5.2) the formula for the degree of the hyperdeterminant:

$$N(1, m-1, m-1) = 2m(m-1), \qquad N(2, m-1, m-1) = 3m(m-1)^2. \tag{5.4}$$

Note that $2 \times m \times m$ is a subboundary format and the first of the formulas (5.4) is consistent with the formula (3.8). It is an easy exercise to deduce the second formula in (5.4) from the general formula (3.6).

For the format $4 \times m \times m$ Theorem 5.2(b) gives us that $\Delta(F_A)$ is equal to the product of $\operatorname{Det}(A)$ and some power R^ν of the Chow form R of $X^\vee_{\text{sing}}$. The value of ν can be obtained by calculation of degrees. By (5.2), the degree of $\Delta(F_A)$ is equal to $4m(m-1)^3$. It follows easily from (3.6) that the degree of $\mathrm{Det}(A)$ is equal to $\frac{2}{3}m(m-1)(m-2)(5m-3)$. On the other hand, the degree of the variety $X^\vee_{\mathrm{sing}}$ is known to be $m^2(m-1)(m+1)/12$; see [1, Chap. 2, formula (5.1)]. Since $X^\vee_{\mathrm{sing}}$ has codimension 4, the degree of its Chow form R as a polynomial in matrix entries of A is four times the degree of $X^\vee_{\mathrm{sing}}$, i.e., is equal to $m^2(m-1)(m+1)/3$. These three expressions imply that the exponent v is equal to 2.

EXAMPLE 5.5. Let $r=3$ and let $V_1=V_2=V_3=\mathbf{C}^2$ be three 2-dimensional vector spaces. Let $V=V_1\otimes V_2\otimes V_3$ be the space of $2\times 2\times 2$ matrices, and e_{ijk} $(i, j, k\in\{0, 1\})$ be its standard basis vectors (matrix units). Let $\nabla\subset V$ be the variety of degenerate matrices. The group $G=GL(V_1)\times GL(V_2)\times GL(V_3)$ acts on the space V, leaving varieties ∇ and ∇_{sing} invariant. It is known and easy to check that G has only seven orbits on V, including $\{0\}$ (and hence six orbits on $P(V)$). The closures of six orbits in $P(V)$ and representatives of these orbits are the following:

dim = 7: $P(V)$ itself; a representative $e_{000}+e_{111}$.

dim = 6: The projectivization $X^\vee$ of ∇; a representative $e_{100}+e_{010}+e_{001}$.

dim = 4: Three varieties

$$P(V_1)\times P(V_2\otimes V_3),\qquad P(V_2)\times P(V_1\otimes V_3),\qquad P(V_1\otimes V_2)\times P(V_3);$$

representatives $e_{010}+e_{001}$, $e_{100}+e_{001}$, $e_{010}+e_{100}$.

dim = 3: The product $P(V_1)\times P(V_2)\times P(V_3)$; a representative e_{000}.

The singular locus $X^\vee_{\mathrm{sing}}$ has three irreducible components, namely the orbit closures of dimension 4 just described. This can be seen by calculation of partial derivatives of the hyperdeterminant of a $2\times 2\times 2$ matrix (given by the formula (1.5)) at all the representatives listed above.

In particular, we see that $X^\vee_{\mathrm{sing}}$ has codimension 3. Hence for a $2\times 2\times 2\times 2$ matrix A we have $\mathrm{Det}(A)=\Delta(F_A)$. This was already known to Schläfli [18, Sect. 19]. The degree of $\mathrm{Det}(A)$ is equal to 24.

For a $3\times 2\times 2\times 2$ matrix A it follows from Theorem 5.2(b) and the obvious symmmetry that for some $v\geqslant 0$

$$\Delta(F_A)=\mathrm{Det}(A)\cdot(R_{12}R_{13}R_{23})^v\,(\mathrm{Im}(\tilde{A})),$$

where R_{ab} is the Chow form of the component $P(V_c)\times P(V_a\otimes V_b)\subset \mathrm{Sing}(X^\vee)$ for $\{a, b, c\}=\{1, 2, 3\}$. The exponent v can be found as in the previous example. By (5.2), the degree of $\Delta(F_A)$ is equal to $3^3\cdot 4=108$. By (3.8), the degree of $\mathrm{Det}(A)$ is equal to $N(2, 1, 1, 1)=2\cdot 3!\cdot 3=36$. Finally, each of the Chow forms $R_{ab}(\mathrm{Im}(\tilde{A}))$ has degree 12 as a polynomial in matrix entries of A (this can be shown, e.g., using the machinery of Section 4; see Remark 4.12(b)). It follows that $\nu = (108 - 36)/3 \cdot 12 = 2$.

Remarks 5.6. (a) It seems very likely that in the general case of Theorem 5.2(b), for any component Z of $X^{\vee}_{\text{sing}}$ the exponent with which R_Z enters $\Delta(F_A)$ is equal to the multiplicity of $X^{\vee}$ along Z (i.e., the degree of the normal cone; see [1, Chap. 2, Sect. 1]). In both Example 5.4 and Example 5.5 the normal cone at a generic point of $X^{\vee}_{\text{sing}}$ can be seen to be a quadratic cone, and the exponent is equal to 2.

(b) It follows from Theorem 5.2 that whenever the hyperdeterminant of a format $(l_1+1) \times \cdots \times (l_r+1)$ exists, we can apply Schläfli's method to matrices of format $2 \times (l_1+1) \times \cdots \times (l_r+1)$ and obtain that $\Delta(F_A)$ is the product of $\text{Det}(A)$ with some extra factors. In particular, this gives a method for calculating the hyperdeterminant of format 2^r by successive computations of discriminants of binary forms. However, the extra factors grow very fast with r.

(c) We conjecture that formats $2 \times m \times m$, $3 \times m \times m$, and $2 \times 2 \times 2 \times 2$ are the only ones for which Schläfli's method gives exactly the hyperdeterminant (i.e., $\Delta(F_A)$ is not identically zero and does not contain extra factors). This is equivalent to the assertion that for any formats other than $m \times m$ and $2 \times 2 \times 2$ the singular locus of ∇ has codimension 2 in the matrix space.

Acknowledgments

Most of the results in this paper were announced in [7]. This work was completed when two of the authors (M.K. and A.Z.) visited the Mathematical Sciences Institute at Cornell University and were partly supported by the U.S. Army Research Office through the MSI. We are very grateful to G.-C. Rota and R. Stanley for their inspiring interest in this work, and to L. Billera and B. Sturmfels for their warm hospitality, moral support, and stimulating discussions.

References

1. E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, "Geometry of Algebraic Curves," Vol. 1, Grundlehren der Math. Wiss., Vol. 267, Springer-Verlag, Berlin/Heidelberg/New York, 1985.
2. J.-M. Bismut and D. S. Freed, The analysis of elliptic families, I, Metrics and connections on determinant bundles, *Comm. Math. Phys.* **106** (1986), 159–176.
3. A. Cayley, On the theory of elimination, *Cambridge Dublin Math. J.* **3** (1848), 116–120; reprinted in "Collected Papers," Vol. 1, No. 59, pp. 370–374, Cambridge Univ. Press, London/New York, 1889.
4. A. Cayley, On the theory of linear transformations, *Cambridge Math. J* **4** (1845), 193–209; reprinted in "Collected Papers," Vol. 1, No. 13, pp. 80–94, Cambridge Univ. Press, London/New York, 1889.

5. A. L. DIXON, The eliminant of three quantics in two independent variables, *Proc. London Math. Soc.* **7** (1908), 49–69.
6. I. M. GELFAND, M. M. KAPRANOV, AND A. V. ZELEVINSKY, Discriminants of polynomials in several variables and triangulations of Newton polytopes, *Algebra i Analiz* **2**, No. 3 (1990), 1–62. [in Russian; English transl., *Leningrad Math. J.* **2**, No. 3 (1991), 449–505.
7. I. M. GELFAND, M. M. KAPRANOV, AND A. V. ZELEVINSKY, Projectively dual varieties and hyperdeterminants, *Soviet Math. Dokl.* **39**, No. 2 (1989), 385–389.
8. P. GRIFFITHS AND J. HARRIS, "Principles of Algebraic Geometry," Wiley–Interscience, New York, 1978.
9. M. JIMBO (Ed.), "Yang–Baxter Equation in Integrable Systems," World Sci. Singapore, 1990.
10. M. M. KAPRANOV, B. STURMFELS, AND A. V. ZELEVINSKY, "Chow Polytopes and General Resultants," *Duke Math. J.* **67**, No. 1 (1992), 189–218.
11. S. KLEIMAN, Enumerative theory of singularities, *in* "Real and Complex Singularities" (Proceedings, Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976)," pp. 297–396, Sijthoff & Noordhoff, Alphen aan den Rijn, 1977.
12. F. KNOP AND G. MENZEL, Duale Varietäten von Fahnenvarietäten, *Comm. Math. Helv.* **62** (1987), 38–61.
13. I. G. MACDONALD, "Symmetric Functions and Hall Polynomials," Clarendon, Oxford, 1979.
14. CH. OKONEK, M. SCHNEIDER, AND H. SPINDLER, "Vector Bundles on Complex Projective Spaces," Progress in Mathematics, Vol. 3, Birkhäuser-Verlag, Boston, 1980.
15. E. PASCAL, "Die Determinanten," Teubner-Verlag, Leipzig, 1900.
16. H. J. RYSER, "Combinatorial Mathematics," Wiley, New York, 1963.
17. D. J. SAUNDERS, "The Geometry of Jet Bundles," London Math. Soc. Lecture Note Series, Vol. 142, Cambridge Univ. Press, Cambridge, 1989.
18. L. SCHLÄFLI, Über die Resultante eines Systemes mehrerer algebraischer Gleichungen, *Denkschr. Kaiserl. Akad. Wiss., Math.-Naturwiss. Klasse* **4** (1852); reprinted in "Gesammelte Abhandlungen," Vol. 2, No. 9, pp. 9–112, Birkhäuser-Verlag, Basel, 1953.
19. B. L. VAN DER WAERDEN, "Einführung in die algebraische Geometrie," Grundlehren der Math. Wiss., Vol. 51, Springer-Verlag, Berlin/Heidelberg/New York, 1973.